Docket No. 13572.105039

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant(s):	Susan Slaugenhaupt et al.		
		Group Art Unit:	1634
Serial No.:	10/041,856		
		Examiner:	Carla Myers
Filed:	January 7, 2002		
		Confirmation No.:	5418

For: Gene For Identifying Individuals With Familial Dysautonomia

DECLARATION OF THE INVENTORS UNDER 37 C.F.R. § 1.131

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

We, Susan A. Slaugenhaupt and James F. Gusella, citizens of the United States of America, residing at 10 Quail Run, Hingham, MA 02043 and 7 Woodstock Drive, Framingham, Massachusetts 01701, respectively, so hereby declare and state that:

1. We are the inventors of the invention that is disclosed and claimed in the above-identified application, Serial No. 10/041,856, filed January 7, 2002, which claims priority to U.S. Application Serial No. 60/260,080, filed January 6, 2001, now abandoned. The claimed invention relates to the genetic mutations that identify individuals with Familial Dysautonomia (FD).

2. We are providing this Declaration to demonstrate that we conceived and reduced to practice at least one embodiment of the claimed invention or an embodiment that would have made the claimed invention obvious at that time prior to December 22, 2000, which is the date of

the document submitted to the USPTO on June 18, 2004 in connection with a Petition Under 37 C.F.R. §1.59(b) to Expunge Information Submitted Under M.P.E.P. §724.02, 8th edition. The date of invention demonstrated herein is also prior to the priority date of January 17, 2001 of the application that published as US 2002/0168656 by Berish Rubin et al., prior to the electronic publication on January 22, 2001, of Slaugenhaupt et al., "Tissue-Specific Expression of a Splicing Mutation in the *IKBKAP* Gene Causes Familial Dysautonomia" published in The American Journal of Human Genetics, volume 68 at pages 598-605, and prior to the electronic publication on January 22, 2001, of Anderson et al. "Familial Dysautonomia is Caused by Mutations of the IKAP Gene" published in The American Journal of Human Genetics, volume 68 at pages 753-758. The date of invention demonstrated herein is also prior to the publication on January 2, 2001 of Gill et al., GenBank Accession No. AF153419, and prior to the publication on February 27, 2001 of Slaugenhaupt et al., GenBank Accession No. AF153419.

3. Prior to December 22, 2000, we submitted a grant application for funding from the National Institutes of Health, for research to identify the FD mutations (Exhibit 1); the grant application documents that we had conceived of the idea to identify FD mutations, and then to detect mutations in prenatal diagnostic testing for FD, as described on page 30 of Exhibit 1.

4. Prior to December 22, 2000, we screened the 168 kilobase candidate region for the FD gene utilizing direct comparison of genomic sequences generated in our laboratory from control and FD individuals and by individual exon amplification, including adjacent intron regions, followed by sequencing in control individuals and individuals with FD. The results of this screening are presented in Exhibit 2. Exhibit 2 at page 3 indicates that, on fragment 25 of the DYS 129 cosmid, an A to G transition was identified 6 base pairs from the intron-exon boundary

in exon 20 and was named GenPoly 130. We had localized the *IKBKAP* gene to at least fragment 25 of the DYS 129 cosmid.

5. Subsequently and also prior to December 22, 2000, we sequenced the intron 20 region from samples from control individuals and individuals with FD and determined that individuals with FD carry a C or G , depending on the strand sequenced, at a position 6 base pairs from the intron/exon boundary of intron 20, while non-carriers consistently have a T or A in that position. Attached hereto as Exhibit 3 is a copy of a notebook page from the laboratory notebook of Sandra Gill, a Senior Research Technologist in the laboratory of Susan A. Slaugenhaupt. The notebook page, which is labeled "GenPoly 130" describes this sequence analysis, the results of which are depicted in the copy of the autoradiograph of the sequencing gel presented in Exhibit 4. The notebook page indicates that all "NC" or non-carriers have a T at the GenPoly 130 position, while the FD samples have a C at that position.

6. From the autoradiograph presented in Exhibit 4, we identified a single DNA base difference between the FD and control sequences, which we called "GenPoly 130," as indicated in Exhibits 3 and 4. This mutation associated with the major haplotype of FD is a T-C change located at base pair 6 of intron 20 of the *IKBKAP* gene. Thus, prior to December 22, 2000, we had identified the major sequence variant associated with FD as a T-C change at base pair 6 of intron 20 of the *IKBKAP* gene. We have reviewed Exhibits 3 and 4, and, although the dates have been removed from these documents, the dates of these documents are prior to December 22, 2000. We also confirm that the gel was run by Sandra Gill, a Senior Research Technologist in the laboratory of and under the direction of Susan Slaugenhaupt in the United States of America.

7. Also prior to December 22, 2000, we screened for and identified a single base pair change in exon 19 unique to four FD chromosomes carrying the minor haplotype. Attached hereto as Exhibit 5 is a copy of a notebook page from the laboratory notebook of James Mull, a Senior Research Technologist in the laboratory of Susan A. Slaugenhaupt. The notebook page is labeled "Intron/Exon Boundary Screening of CG5" and "GenPoly 190." The notebook page describes the sequence analysis depicted on the copy of the autoradiogram of the sequence gel provided in Exhibit 6. Exhibit 6 identifies a G to C mutation which is 73 base pairs from the intron/exon boundary of exon 19.

8. From Exhibit 6, we identified a single DNA base difference between sequences from FD and control samples, which sequence variant we called "GenPoly 190". This mutation associated with the minor haplotype of FD is a G-C change at base pair 2397 of the *IKBKAP* coding sequence or base pair 79 of exon 19 of the *IKBKAP* gene. We have reviewed Exhibits 5 and 6 and confirm that, although the dates have been removed from these documents, the dates of these documents are prior to December 22, 2000. We also confirm that all acts relied upon in Exhibits 5 and 6 were carried out by one or more of us, or by a technician in the laboratory under our direction.

9. Having identified the mutations associated with the major and minor haplotypes of Familial Dysautonomia, we had conceived of methods, oligonucleotides, and kits for detecting the FD mutations in samples using routine nucleic acid diagnostic technology. Once we had the mutations, it was absolutely routine to generate oligonucleotide probes and primers to be used in methods to detect mutations in potential carriers of FD.

10. We declare further that all statements made in this Declaration of our own knowledge are true and that all statements made on information and belief are believed to be true

and further that these statements are made with the knowledge that willful false statements and the like are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Date: 2-28-07

Susan A. Slaugenhaupt

Date: 2/28/07

James F.Gusella

EXHIBIT 1

DESCRIPTION: State the application's broad, long-term objectives and specific aims, making reference to the health relatedness of the project. Describe concisely the research design and methods for achieving these goals. Avoid summaries of past accomplishments and the use of the first person. This description is meant to serve as a succinct and accurate description of the proposed work when separated from the application. If the application is funded, this description, as is, will become public information. Therefore, do not include proprietary/confidential information. **DO NOT EXCEED THE SPACE PROVIDED.**

Familial dysautonomia (FD; Riley-Day syndrome) is the best known and most frequent of a group of congenital sensory neuropathies characterized by widespread sensory and variable autonomic dysfunction. First described in 1949, FD is a devastating disorder that involves progressive neuronal degeneration with a broad impact on the operation of many of the body's systems leading to a vastly reduced quality of life and premature death. Affected individuals demonstrate lack of overflow tears, impaired temperature and pain sensation, and autonomic dysfunction, including labile blood pressure and uncoordinated swallowing. Despite recent advances in the management of FD, the disorder is inevitably fatal with only 50% of patients reaching 30 years of age.

FD is due to a recessive genetic defect with a remarkably high carrier frequency in Ashkenazi Jews of 1 in 30, rivaling the gene frequencies of more widely recognized disorders such as Tay-Sachs disease and cystic fibrosis. FD's genetic characteristics make it ideally suited for molecular investigation. We have used genetic linkage to map the defective gene to chromosome 9q31 and have determined that its ethnic bias is due to a founder effect, with most disease alleles sharing a common ancestral mutation. This has provided a strategy for narrowing the location of the disease gene to a small stretch of genomic DNA. We are now poised to identify the nature of the genetic defect in FD and to characterize its mode of pathogenesis. The product of this work will be a knowledge of the cause of FD and its relationship to other sensory neuropathies, the ability to screen for carriers of the disorder, and a beginning to the exploration of the normal and abnormal function of the DYS gene in human and animal model systems. In the long-term we will contribute both to the fundamental understanding of development and maintenance of the sensory and autonomic nervous systems, and to the hope of FD patients for an effective treatment preventing progression of this devastating disorder.

PERFORMANCE SITE(S) (*organization, city, state*)

Massachusetts General Hospital, Boston MA 02114

KEY PERSONNEL. See instructions on Page 11. *Use continuation pages as needed* to provide the required information in the format shown below.

Name	Organization	Role on Project
James F. Gusella, Ph.D.	Massachusetts General Hospital	Principal Investigator
Susan A. Slaugenhaupt, Ph.D.	Massachusetts General Hospital	Co-investigator
Vanessa Wheeler, Ph.D.(pend.)	Massachusetts General Hospital	Research Associate

Type the name of the principal investigator/program director at the top of each printed page and each continuation page. (For type specifications, see instructions on page 6.)

RESEARCH GRANT

TABLE OF CONTENTS

Research Plan

*Type density and type size of the entire application must conform to limits provided in instructions on page 6.

Appendix (*Five collated sets. No page numbering necessary for Appendix.*) [X] Check if Appendix is included

Number of publications and manuscripts accepted or submitted for publication (*not to exceed 10*) 10

Other items (*list*):

PERSONNEL

Dr. James F. Gusella, Ph.D., P.I. is a molecular geneticist with a well-established track record in the investigation of human genetic disorders. His laboratory has mapped of numerous neurogenetic disease genes, including the defect in familial dysautonomia, and has been responsible for isolation and characterization of the Huntington disease gene, the merlin tumor suppressor of neurofibromatosis 2 and the Batten disease gene. Dr. Gusella will devote 20% effort to the cloning and characterization of the familial dysautonomia gene and will be responsible for the overall direction and implementation of the project, planning and execution of the studies, and the supervision of the research staff.

Susan Slaugenhaupt, Ph.D., co-inv., is a faculty member of the Molecular Neurogenetics Unit who has been involved in the familial dysautonomia project from its early days. She has extensive experience in genetic analysis, having trained both in the statistical aspects of human genetics and in the molecular biological techniques of positional cloning. Dr. Slaugenhaupt has been responsible for day-to-day direction of the technical staff on this project and for driving the completion of the physical and genetic mapping of the DYS gene. In the spring of 1997, she will be moving into new laboratory space being developed for the Molecular Neurogenetics Unit as part of the Harvard Institute of Human Genetics where this project will be carried out. Dr. Slaugenhaupt will devote 50% effort to the project, aiding Dr. Gusella in the planning and implementation of the studies, and providing day-to day supervision of the research associate and other staff.

Vanessa Wheeler, Research Associate, will be joining the lab in the spring of 1997 after completing the requirements for her Ph.D. at St Mary's Hospital in London. She has already received extensive training in basic molecular biology, and will receive specific experience in positional cloning techniques as part of this project. Vanessa will devote 100% effort to the cloning and characterization of the DYS gene.

Blake Liebert, M.Sc., is a senior research technologist who has been a mainstay of the familial dysautonomia group since discovery of genetic linkage for the DYS gene. He is experienced in all of the molecular biological techniques used in this labor intensive project and will devote 100% effort to it. Blake is also responsible for ordering and maintaining the inventory of supplies for this project.

Jim Mull, B.Sc., is a research technologist who has worked with Dr. Slaugenhaupt for the past two years on DNA marker analysis, cloning and DNA sequencing in the DYS candidate region. He will devote 50% effort to these tasks as part of this project. He will also be responsible for maintaining all clone stocks for the present project.

EQUIPMENT

This project will be carried out in new laboratory space being made available to Dr. Gusella, Dr. Slaugenhaupt and the MGH Molecular Neurogenetics Unit as part of the new Harvard Institute of Human Genetics. As it is up to individual investigators to provide the basic equipment for this new space, we are requesting several basic items that will be necessary for successful pursuit of this project. These include: a 4 ft. Sterilgard SG-4005 laminar flow hood ($5600), a laboratory refrigerator ($500), a -20°C laboratory freezer ($500), and a -80°C laboratory freezer ($6,000).

No equipment is requested for subsequent years. The total of the equipment requested is $12,600.

SUPPLIES

The supplies required for these studies can be classified into 6 categories:

Tissue culture supplies (culture media, fetal calf serum, CO_2, plasticware etc.)	$ 4,000
Molecular biological reagents (restriction enzymes, DNA ligase, DNA polymerase, RNA polymerases, agarose, nylon membrane, hybridization reagents, etc.)	$12,000
PCR amplification and DNA sequencing reagents (sequencing kits, oligonucleotides, Pfu and Taq polymerases etc.)	$ 9,000
Bacteriological supplies (media, agar, bacterial plasticware, Xgal, antibiotics, centrifuge bottles, etc.)	$ 4,000
Radiochemical supplies (α and γ ^{32}P nucleotides, ^{125}I antibodies, ^{35}S methionine, X-ray film, etc.)	$ 8,000
General laboratory chemicals and disposable supplies (Buffers, salts, solvents, disposable plasticware, pipettes, pipette tips, disposable gloves, polaroid film, etc.)	$ 8,000
Total	$45,000

This proposed project entails the efforts of ~3.2 investigators in the laboratory performing a combination of molecular studies at the DNA, RNA and protein levels. Positional cloning projects typically require higher supply spending than other more limited molecular biological studies because of the large numbers of clones being isolated, mapped and analyzed. Based on extensive experience in this type of project, we estimate the cost of the supplies for the first year to be $45,000, or approximately $14,000/investigator.

TRAVEL

$2,400.00 is requested for Drs Gusella, Slaugenhaupt and Wheeler to attend a national meeting on genetics or molecular neurobiology, such as the American Society of Human Genetics or the Society for Neuroscience annual meetings.

OTHER

Radiation Monitoring and Waste Disposal	$ 3,500
Project share of service contracts on shared major equipment	$ 3,000
VAX Computer costs, lab notebooks, publication costs, overnight mail, etc.	$ 3,000
Use of institutional core research services	$ 2,500
Total	$12,000

Fringe benefits rates are 26.47% and 21.68% professional and non-professional staff, respectively.

BIOGRAPHICAL SKETCH

Give the following information for the key personnel and consultants and collaborators. Begin with the principal investigator/program director. Photocopy this page for each person.

NAME James F. Gusella, Ph.D. POSITION TITLE Professor of Genetics

EDUCATION *(Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)*

INSTITUTION AND LOCATION	DEGREE	YEAR CONFERRED	FIELD OF STUDY
University of Ottawa, Ottawa, Canada	B.Sc.	1974	Honours Biology
University of Toronto, Toronto, Canada	M.Sc.	1976	Medical Biophysics
Massachusetts Institute of Technology	Ph.D.	1980	Genetics

RESEARCH AND/OR PROFESSIONAL EXPERIENCE: Concluding with present position, list in chronological order previous employment, experience, and honors. Key personnel include the principal investigator and any other individuals who participate in the scientific development or execution of the project. Key personnel typically will include all individuals with doctoral or other professional degrees, but in some projects will include individuals at the masters or baccalaureate level provided they contribute in a substantive way to the scientific development or execution of the project. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. DO NOT EXCEED TWO PAGES.

PROFESSIONAL APPOINTMENTS:

1980-1984 Instructor in Neurology (Genetics), Harvard Medical School
1983-1984 Assistant Geneticist, Neurology Service, MGH
1984-1987 Assistant Professor of Genetics, Harvard Medical School
1984-1992 Associate Geneticist, Neurology Service, MGH
1984- Director, Molecular Neurogenetics Unit, Massachusetts General Hospital
1987- Associate Professor of Genetics, Harvard Medical School
1988-1992 Mallinckrodt Associate Professor of Genetics, HMS
1992- Geneticist, Neurology Service, MGH
1992- Professor of Genetics, HMS
1993- Bullard Professor of Neurogenetics, HMS

HONORS AND AWARDS (selected):

1986 Jordi Folch Memorial Award, American Society for Neurochemistry
1987 Wadsworth Award, New York State Department of Health
1987 A. Cressy Morrison Award in the Natural Sciences, New York Academy of Science
1987 Metropolitan Life Foundation Award for Medical Research
1988 Bennett Award Lecture, American Neurological Association
1993 National Health Council Award for Medical Research
1994 The Herb Boyer Lecture, University of California at San Francisco
1994 J. Allyn Taylor International Prize in Medicine

ORIGINAL PUBLICATIONS (selected from 350)

Gusella J, Wexler N, Conneally P, Naylor S, Anderson M, Tanzi R, Watkins P, Ottina K, Wallace M, Sakaguchi A, Young A, Shoulson I, Bonilla E, Martin J. A polymorphic DNA marker genetically linked to Huntington's Disease. Nature 1983; 306: 234-238.

St George-Hyslop P, Haines J, Farrer L, Polinsky R, Van Broeckhoven C, Goate A, Crapper McLachlan D, Orr H, Bruni A, Sorbi S, Rainero I, Foncin J, Pollen D, Cantu J, Tupler R, Voskresenskaya N, Mayeux R, Growdon J, Nee L, Backhovens H, Martin J, Rossor M, Owen M, Mullan M, Percy M, Karlinsky H, Rich S, Heston L, Montes M, Mortilla M, Nacmias N, Vaula G, Gusella J, Hardy J. Genetic linkage studies suggest that Alzheimer's disease is not a single homogeneous entity. Nature 1990; 347: 194-197.

Fontaine B, Khurana T, Hoffman E, Bruns G, Haines J, Trofatter J, Hanson M, Rich J, McFarlane H, Yacek D, Gusella J, Brown R. Hyperkalemic periodic paralysis and the adult muscle sodium channel a-subunit. Science 1990; 250: 1000-1002.

MacDonald M, Novelletto A, Lin C, Tagle D, Barnes G, Bates G, Taylor S, Allitto B, Altherr M, Myers R, Lehrach H, Collins F, Wasmuth J, Frontali M, Gusella J. The Huntington's disease candidate region exhibits many different haplotypes. Nature Genet 1992: 1: 99-103.

McClatchey A, Van den Berg P, Pericak-Vance M, Raskind W, Verellen C, McKenna-Yasek D, Rao K, Haines J, Bird T, Brown R, Gusella J. Temperature-sensitive mutations in the III-IV cytoplasmic loop region of the skeletal muscle sodium channel gene in paramyotonia congenita. Cell 1992; 68: 769-774.

McClatchey A, McKenna-Yasek D, Cros D, Worthen H, Kuncl R, DeSilva S, Cornblath D, Gusella J, Brown, R. Novel mutations in families with unusual and variable disorders of the skeletal muscle sodium channel. Nature Genetics 1992; 148-152.

The I, Murphy A, Hannigan G, Jacoby L, Menon A, Gusella J, Bernards A. Neurofibromatosis type 1 gene mutations in neuroblastoma. Nature Genetics 1993; 3: 62-66.

Bernards A, Snijders A, Hannigan G, Murthy A, Gusella J. Mouse neurofibromatosis type 1 cDNA sequence reveals high degree of conservation of both coding and non-coding mRNA segments. Hum Mol Genet 1993; 2: 645-650.

Rosen D, Siddique T, Patterson D, Figliewicz D, Sapp P, Hontatit A, Donaldson D, Goto J, O'Regan J, Deng H, Rahmani Z, Krizus A, McKenna-Yasek D, Cayabyab A, Gaston S, Tanzi R, Halperin J, Herzfeldt B, Van den Berg R, Hung W, Bird T, Deng G, Mulder D, Smyth C, Laingi N, Soriano E, Pericak-Vance M, Haines J, Rouleau G, Gusella J, Horvitz H, Brown R. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 1993; 362: 59-62.

The Huntington's Disease Collaborative Research Group A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 1993; 72: 971-983.

Trofatter J, MacCollin M, Rutter J, Murrell J, Duyao M, Parry D, Eldridge R, Kley N, Menon A, Pulaski K, Haase V, Ambrose C, Munroe D, Bove C, Haines J, Martuza R, MacDonald M, Seizinger B, Short M, Buckler A, Gusella J. A novel *moesin-, ezrin-, radixin*-like gene is a candidate for the neurofibromatosis 2 tumor suppressor. Cell 1993; 72: 791-800.

Tanzi R, Petrukhin K, Chernov I, Pallequer J, Wasco W, Ross B, Romano D, Parano E, Pavone L, Brzustowicz L, Devoto M, Peppercorn J, Bush A, Sternlieb I, Perastu M, Gusella J, Evgrafov O, Penchaszadeh G, Honig B, Edelman I, Soares M, Scheinberg I, Gilliam T. The Wilson disease gene is a copper transporting ATPase with homology to the Menkes disease gene. Nature Genet 1993; 5: 344-350.

Blumenfeld A, Slaugenhaupt S, Axelrod F, Lucente DE, Maayan C, Liebert C, Ozelius L, Trofatter J, Haines J, Breakefield X, Gusella J. Localization of the gene for familial dysautonomia on chromosome 9 and definition of DNA markers for genetic diagnosis. Nature Genet 1993; 4:160-164.

Ambrose C, Duyao M, Barnes G, Bates G, Lin C, Srinidhi J, Baxendale S, Hummerich H, Lehrach H, Altherr M, Wasmuth J, Buckler A, Church D, Housman D, Berks M, Micklem G, Durbin R, Dodge A, Read A, Gusella J, MacDonald M. Structure and expression of the Huntington's disease gene: Evidence against simple inactivation due to an expanded CAG repeat. Somat Cell Mol Genet; 1994; 20, 27-38.

MacCollin M, Ramesh V, Jacoby L, Louis D, Rubio M, Pulaski K, Trofatter J, Short M, Bove C, Eldridge R, Parry D, Gusella J. Mutational analysis of patients with neurofibromatosis 2. Am J Hum Genet 1994; 55: 314-320.

Persichetti F, Srinidhi J, Kanaley L, Ge P, Myers R, D'Arrigo K, Barnes G, MacDonald M, Vonsattel J, Gusella J, Bird E. Huntington's disease CAG trinucleotide repeats in pathologically confirmed post-mortem brains. Neurobiol Dis 1994; 1: 159-166.

Duyao M, Auerbach A, Ryan A, Persichetti F, Barnes G, McNeil S, Ge P, Vonsattel J, Gusella J, Joyner A, MacDonald M. Inactivation of the mouse Huntington's disease gene homolog *Hdh*. Science 1995; 269, 407-410.

Persichetti F, Ambrose C, Ge P, McNeil S, Srinidhi J, Anderson M, Jenkins B, Barnes G, Duyao M, Kanaley L, Wexler N, Myers R, Bird E, Vonsattel J, MacDonald M, Gusella J. Normal and expanded Huntington's disease gene alleles produce distinguishable proteins due to translation across the CAG repeat. Mol Med 1995; 1: 374-83.

The International Batten Disease Consortium: Group 1; Lerner TJ, Boustany RMN, Anderson JW, D'Arigo KL, Schlumpf K, Buckler AJ, Gusella JF, Haines JL. Group 2; Kremmidiotis G, Lensink IL, Sutherland GR, Callen DF. Group 3; Taschner PEM, de Vos N, van Ommen GJB, Breuning MH. Group 4; Doggett NA, Meincke LJ, Liu ZY, Goodwin LA, Tesmer JG. Group 5; Mitchison HM, O'Rawe AM, Munroe PB, Jarvela IE, Gardiner RM, Mole SE. Isolation of a novel gene underlying Batten disease, *CLN3*. Cell 1995; 82: 949-957.

BIOGRAPHICAL SKETCH

rovide the following information for the key personnel in the order listed on Form Page 2.
Photocopy this page or follow this format for each person.

NAME	POSITION TITLE
Slaugenhaupt, Susan A.	Assistant in Genetics (Neurology, MGH) Instructor in Neurology (HMS)

EDUCATION/TRAINING (*Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.*)

INSTITUTION AND LOCATION	DEGREE (*if applicable*)	YEAR(s)	FIELD OF STUDY
Eckerd College, St. Petersburg, FL	B.S.	1985	Biology
University of Pittsburgh, Pittsburgh, PA	M.S.	1988	Human Genetics
University of Pittsburgh, Pittsburgh, PA	Ph.D.	1991	Human Genetics

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. If the list of publications in the last three years exceeds two pages, select the most pertinent publications. **DO NOT EXCEED TWO PAGES.**

PROFESSIONAL APPOINTMENTS:

1991 - 1995	Research Fellow in Neurology, Molecular Neurogenetics Unit, Massachusetts General Hospital
1991 - 1995	Research Fellow in Neurology, Harvard Medical School
1995 -	Assistant in Genetics (Neurology), Massachusetts General Hospital
1995 -	Instructor in Neurology, Harvard Medical School

HONORS AND AWARDS:

1981 - 1990	Wilbur C. Stauble Trust Scholarship, Western Pacific Industries
1982 - 1985	Honor Scholarship, Eckerd College
1985	National Dean's List
1985	Who's Who Among Students in American Colleges and Universities
1986	Travel Award - Seventh International Congress of Human Genetics, Berlin, 1988, 1990
	Awardee - University of Pittsburgh Honors Convocation

ORIGINAL PUBLICATIONS (selected from 38)

Warren AC, Chakravarti A, Wong C, Slaugenhaupt SA, Halloran S, Watkins PC, Metaxotou C,Antonarakis SE. 1987. Evidence for reduced recombination on the nondisjoined chromosomes 21 in Down syndrome. Science. 237:652.

Chakravarti A and Slaugenhaupt SA. 1987. Methods for studying recombination on chromosomes that undergo nondisjunction. Genomics. 1:35.

Morizot DC, Slaugenhaupt SA, Kallman KD, Chakravarti A. 1991. Genetic linkage map of fishes of the genus *Xiphophorus* (Teleostei: Peociliidae). Genetics 127:399.

Petersen MB, Slaugenhaupt SA, Lewis JG, Warren AC, Chakravarti A, Antonarakis SE. 1991. A genetic linkage map of 27 markers on human chromosome 21. Genomics 9:407.

Lupski, JR, Montes de Oca-Luna R, Slaugenhaupt SA, Pentao L, Guzzetta V, Trask BJ, Saucedo-Cardesas O, Barker D, Killian JM, Garcia CA, Chakravarti A, Patel PI. 1991. DNA Duplication Associated with Charcot-Marie-Tooth Disease Type 1A. Cell. 66:1.

Kwiatkowski DJ, Povey S, Armour J, Attwood J, Furlong RA, Goudie DR, Haines JL, Pericak-Vance MA, Slaugenhaupt SA, Vergnaud G, Warnich L, Yuille MAR. 1992. A Comprehensive Genetic Linkage Map of the Human Genome: Chromosome 9. Science. 258:67-86.

Blumenfeld A, Axelrod FB, Trofatter JA, Maayan C, Lucente DE, Slaugenhaupt SA, Liebert CB, Ozelius LJ, Haines JL, Breakefield XO, Gusella JF. 1993. Exclusion of familial dysautonomia from more than 60% of the genome. J. Med. Genet. 30:47-52.

McClatchey AI, Cannon SC, Slaugenhaupt SA, Gusella JF. 1993. The cloning and expression of a sodium channel b1-subunit from human brain. Hum. Mol. Genet. 2(6):745-749.

Blumenfeld A, Slaugenhaupt SA, Axelrod FB, Lucente DE, Maayan C, Liebert CB, Ozelius LJ, Trofatter JA, Haines JL, Breakefield XO, Gusella JF. 1993. Localization of the gene for familial dysautonomia on chromosome 9 and definition of DNA markers for genetic diagnosis. Nature Genetics. 4:160-164.

Angrist M, Kauffman E, Slaugenhaupt SA, Matise TC, Puffenburger EG, Washington SS, Lipson A, Cass DT, Reyna T, Weeks DE, Sieber W, and Chakravarti A. 1993. A gene for Hirschsprung disease (megacolon) in the pericentromeric region of human chromosome 10. Nature Genetics. 4:451-356.

Velinov M, Slaugenhaupt SA, Stoilov I, Scott CI, Gusella JF, and Tsipouras P. 1994. The achondroplasia gene maps to the telomeric region of chromosome 4p. Nature Genetics. 6(3):314-317.

Puffenberger EG, Kauffman ER, Bolk S, Matise TC, Washington SS, Angrist M, Weissenbach J, Garver KL, Mascari M, Ladda R, Slaugenhaupt SA, Chakravarti AC. 1994. Identity-by-descent and association mapping of a recessive gene for Hirschsprung disease on human chromosome 13q22. Hum. Mol. Genet. 3(8):1217-1225.

Slaugenhaupt SA, Blumenfeld A, Liebert CB, Mull J, Lucente DE, Monahan M, Breakefield XO, Axelrod FB, Maayan C, Parada L, and Gusella, JF. 1995. The human gene for neurotrophic tyrosine kinase receptor type 2 is located on chromosome 9 but is not the familial dysautonomia gene. Genomics. 25:730-732.

Slaugenhaupt SA, Roca AL, Lievert, CB, Altherr MR, Gusella, JR, and Reppert, SM. 1995. Mapping of the gene for the Mel_{1a}-Melatonin Receptor to Human Chromosome 4 (*MTNR1A*) and Mouse Chromosome 8 (*Mtnr1a*). Genomics. 27:355-357.

Sahin M, Slaugenhaupt SA, Gusella JF, and Hockfield, S. 1995. A Protein Tyrosine Phosphatase, PTPH1, is expressed in a segment-specific pattern in the rat diencephalon. Proc. Natl. Acad. Sci. USA. 92:7859-7863.

Oddoux C, Reich E, Axelrod F, Maayan C, Blumenfeld A, Slaugenhaupt SA, Gusella JF, Ostrer H. 1995. Prenatal diagnostic testing using for familial dysautonomia using linked genetic markers. Prenatal Diagnosis 15:817-826.

Eng CM, Slaugenhaupt SA, Blumenfeld A, Axelrod FB, Gusella JF, and Desnick RJ. 1995. Prenatal diagnosis of familial dysautonomia by analysis of linked CA-repeat polymorphisms on chromosome 9q31-q33. Am. J. Med. Genet.59:349-355.

Reppert SM, Godson C, Mahle CD, Weaver DR, Slaugenhaupt SA, Gusella JF. 1995. Molecular characterization of a novel melatonin receptor expressed in human retina and brain: the Mel_{1b}-melatonin receptor. Proc. Natl. Acad. Sci. USA. 92:8734-8738.

Blumenfeld A, Lucente DE, Trofatter JA, Lerner T, Slaugenhaupt SA, Liebert CB, Monahan M, Haines JL, Gusella JF, Breakefield XO, and Parysek LM. 1995. Peripherin Gene is linked to Keratin 18 Gene on Human Chromosome 12. Som. Cell. Molec. Genet. 21(1):83-88.

Slaugenhaupt SA, Liebert CB, Mull J, MacCormack K, LeBel A, Buckler A, McCormick MK, Ekong R, Gusella J. 1996. The gene for tumor necrosis factor receptor associated factor 2 maps to human chromosome 9q34. (in preparation)

Slaugenhaupt SA, Liebert CB, Mull J, MacCormack K, LeBel A, Buckler A, McCormick MK, Blumenfeld A, Gusella J. 1996. Evaluation of PTPH1, a protein tyrosine phosphatase, as a candidate gene for Familial dysautonomia. (in preparation)

Slaugenhaupt SA, Liebert CB, Mull J, MacCormack K, LeBel A, Buckler A, McCormick MK, Ekong R, Gusella J. 1996. Localization and genomic structure of the human PTPA gene to chromosome 9q34. (in preparation)

BIOGRAPHICAL SKETCH

Provide the following information for the key personnel in the order listed on Form Page 2.
Photocopy this page or follow this format for each person.

NAME	POSITION TITLE
Wheeler, Vanessa C.	Research Associate

EDUCATION/TRAINING (*Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.*)

INSTITUTION AND LOCATION	DEGREE (*if applicable*)	YEAR(s)	FIELD OF STUDY
Sidney Sussex College, University of Cambridge	BA	1992	Natural Sciences
St. Mary's Hospital Medical School, Imperial College, University of London	Ph.D.	1996 (expected)	Molecular Biology

RESEARCH AND PROFESSIONAL EXPERIENCE: Concluding with present position, list, in chronological order, previous employment, experience, and honors. Include present membership on any Federal Government public advisory committee. List, in chronological order, the titles, all authors, and complete references to all publications during the past three years and to representative earlier publications pertinent to this application. If the list of publications in the last three years exceeds two pages, select the most pertinent publications. **DO NOT EXCEED TWO PAGES.**

RESEARCH EXPERIENCE

1991 Leukaemia Research Fund summer studentship:
Molecular detection of *bcr-abl* translocations in patients with chronic myeloid leukaemia

1992 BA Part II project:
Expression of novel insecticidal endotoxin genes in *Bacillus thuringiensis*

1992 - Ph.D. Project
Investigation of the potential of the mitochondrial genome as a means of delivery and expression of genes for gene therapy.

ORIGINAL PUBLICATIONS

Crickmore N, Wheeler VC, and Ellar DJ. 1994. Use of an operon fusion to induce expression and crystallisation of a *Bacillus thuringiensis* -endotoxin encoded by a cryptic gene. *Mol. Gen. Genet.* 242:365-368.

Wheeler VC, Prodromou C, Pearl L, Williamson R, and Coutelle C. 1994. PCR synthesis, cloning and expression of human ornithine transcarbamylase compatible with mitochondrial and universal codon usage. Second meeting of the European Working Group on Human Gene Transfer and Therapy. *Gene Therapy 1* (suppl. 2) A98:25.

Wheeler VC, and Coutelle C. 1995. Nondegradative *in vitro* labeling of plasmid DNA. *Anal. Biochem.* 225:374-376.

Wheeler VC, Prodromou C, Pearl L, Williamson R, and Coutelle C. 1996. Synthesis of a modified gene coding for human ornithine transcarbamylase for expression in mammalian mitochondrial and universal translation systems: a novel approach towards correction of a genetic defect. *Gene* (accepted).

OTHER SUPPORT:

James F. Gusella, Ph.D.

2 PO1 NS16367-16 : (Gusella) 07/01/80-06/30/00 25%
NIH/NINDS $159,337 (Proj.1)
$ 33,000 (Core A)

Huntington's Disease Center Without Walls (Director, JF Gusella)
This multidisciplinary program project grant funds the work of 5 laboratories aimed at genetic, biochemical and anatomical analyses of HD.

Project 1: Consequences of expanded CAG in HD (PI: JF Gusella)
This project is aimed at the genetics of CAG expansion in HD and the biochemical and cell biological characterization of normal and mutant huntingtin.

Core A: Administration
This core provides administrative support for the program project.

2 PO1 NS24279-10A1: (Breakefield) 06/01/95-05/31/00 10%
NIH/NINDS $120,825 (Project 1)

Molecular Genetics of Inherited Neurologic Diseases
Project 1: Characterization of Merlin Expression.
This grant is aimed at the characterization of the merlin NF2 tumor suppressor.

5 RO1 CA57683-04: (Gusella) 08/01/92-07/31/97 10%
NCI $173,468

Toward a Molecular Classification of Human Astrocytomas
This grant is aimed at molecular analysis of astrocytomas. In the competing renewal application it will be assumed by the Co-PI, Dr. David Louis.

2 RO1 NS22224-11A1: (Gusella) 02/29/96-01/31/00 10%
NIH NINDS $186,210

Genetic and Molecular Studies of NF1
This grant is aimed at mutational and functional analyses of neurofibromin.

US ARMY MEDICAL RESEARCH ACQUISITION ACTIVITY DAMD17-93-V-3017: (Gusella)
09/22/93-09/21/96 10%
$210,238
Mutation Analysis of the NF2 Gene
This grant, aimed at mutational analysis of the NF2 gene will end in September 96 and is not renewable.

5 RO1 HG00169-12: (Gusella) 06/15/90-06-30-97 20%

NIH-HG $314,943

Extracting genes from cloned sequenced DNA

This grant is aimed at comparing exon trapping, direct cDNA selection and sequence-based exon prediction as methods of finding genes. The knowledge gained is directly applicable to the current project, but the grant itself is aimed at a cloned, sequenced region of chromosome 4. It will not be renewed.

Bristol Myers Squibb Academic Alliance 0%

(Gusella) 10/01/95-09/30/97

$434,783

This departmental grant provides supplemental support for the investigators of the Molecular Neurogenetics Unit and the Genetics and Aging Unit, with the majority of the funding being aimed at Alzheimer disease. None of the funding is applicable to familial dysautonomia.

Dysautonomia Foundation 0%

(Gusella) 01/01/96-12/31/96

$357,999

Linked and Flanking Markers for Familial Dysautonomia.

This is a contract that has been aimed at developing an accurate prenatal test for familial dysautonomia based on linkage studies. It supported the progress outlined in the current proposal but will terminate at the end of this year. It is our hope to continue the project to identification and characterization of the familial dysautonomia defect with support from this NIH proposal.

Muscular Dystrophy Association: 0%

(Gusella) 07/01/94-06/30/96

$50,000

Linkage and cloning of familial spastic paraperesis

This seed grant supports Unit linkage studies of familial spastic paraperesis kindreds.

NO OVERLAP

OTHER SUPPORT

Susan A. Slaugenhaupt, Ph.D.

NONE

NO OVERLAP

Vanessa Wheeler, Ph.D.

NONE

NO OVERLAP

RESOURCES

FACILITIES: Specify the facilities to be used for the conduct of the proposed research. Indicate the performance sites and describe capacities, pertinent capabilities, relative proximity, and extent of availability to the project. Under "Other," identify support services such as machine shop, electronics shop, and specify the extent to which they will be available to the project. Use continuation pages if necessary.

Laboratory: The work will be conducted in 2,000 sq.ft. of newly constructed laboratory space in the new Harvard Institute of Human Genetics at the Harvard Institute of Medical Sciences Building. The space will be owned and operated by the Massachusetts General Hospital and occupied by the MGH Molecular Neurogenetics Unit as their participation in this new interinstitutional cooperative venture in human genetics.

Clinical: Not applicable

Animal: The Harvard Institute of Medical Sciences Building is equipped with a modern barrier facility for animal research.

Computer: The laboratory is tied into the Harvard and MGH computer networks with access to the Internet and accounts on the MGH HELIX VAX computer for data management and analysis. We also have access to several SPARCstations and PCs running a variety of databasing, analytical, and word processing software.

Office: There is approximately 500 sq.ft. of office space available for the PI and co-investigator in the new research facility. The offices both enter directly into the lab space.

Other: .

MAJOR EQUIPMENT: List the most important equipment items already available for this project, noting the location and pertinent capabilities of each.

The Molecular Neurogenetics Unit at the MGH is fully equipped for molecular biological research. However, with the opening of new space at the Harvard Campus we are requesting a few items of commonly used equipment are requested to provide the basic resources in the newly developed lab space to effectively carry out this project.

a. Specific Aims

Familial dysautonomia (FD; Riley-Day syndrome) is the best known and most frequent of a group of congenital sensory neuropathies characterized by widespread sensory and variable autonomic dysfunction. First described in 1949, FD is a devastating disorder that involves progressive neuronal degeneration with a broad impact on the operation of many of the body's systems leading to a vastly reduced quality of life and premature death. FD is due to a recessive genetic defect with a remarkably high carrier frequency in Ashkenazi Jews of 1 in 30, rivaling the gene frequencies of more widely recognized disorders such as Tay-Sachs disease and cystic fibrosis.

Unlike the sensory motor neuropathies, little is known concerning the causes of neuronal dysfunction and loss in the sensory autonomic neuropathies. However, FD's genetic characteristics make it ideally suited for molecular investigation. We have used genetic linkage to map the defective gene to chromosome 9q31 and determined that its ethnic bias is due to a founder effect, with most disease alleles sharing a common ancestral mutation. This has provided a strategy for narrowing the location of the disease gene to a small stretch of genomic DNA. We are now poised to identify the FD gene by positional cloning and to characterize the nature of the genetic defect.

Specifically, we aim:

1. To complete the physical and genetic definition of the FD candidate region. The disease gene has been confined to a small segment of 9q31 much of which is already cloned in a cosmid and YAC contig. We will complete the cloning of a contig spanning the entire candidate region, along the way generating informative genetic markers to firmly position the sites of recombination events that define the minimum FD candidate region. Cloned segments will be used as probes to directly compare FD and normal chromosomes for evidence of deletions or other rearrangements.
2. To saturate the candidate region with transcripts. We will use exon trapping to generate coding sequence probes for screening of cDNA libraries. cDNA clones will be characterized by sequence, expression pattern and extent of genomic coverage. If any significant regions of the candidate region remain unrepresented by cDNAs, we will use direct cDNA selection and sequence-based exon prediction as alternative gene finding strategies to fill the gaps.
3. To identify the genetic defect and characterize the mutation spectrum. We will prioritize the candidate cDNAs by their location, by any apparent function implied by their sequence and by their expression pattern in FD and controls. Direct sequencing of RT-PCR fragments will be used as the most efficient comprehensive method of scanning the genes from several FD haplotypes for the expected "loss-of function" mutations that would identify the disease gene. Alternative strategies will be explored in the unlikely event that no coding sequence mutation is found. When the disease gene is identified, mutations representing the different FD haplotypes will be determined, and other sensory neuropathies will be examined for defects in this gene or any obvious family members.
4. To characterize the protein product and define the mode of pathogenesis. The strategies to be used for this aim are not completely predictable as they are dependent on the precise nature of the FD protein, but likely possibilities include genotype:phenotype correlations, antibody preparation for examining the protein, defining the relationship to related genes in man and other organisms, exploration of the gene's expression in established experimental paradigms, development of a tissue culture system with an assayable phenotype, generation of knock-out mice as a disease model and many others. The precise approaches to be used will be decided upon based on the nature of the FD gene and its mutations.

b. Background and Significance

Familial dysautonomia (FD; hereditary sensory neuropathy type III) with its extensive sensory and variable autonomic dysfunction, is the most common and widely recognized of the congenital sensory neuropathies (1). FD has been the most intensively studied of any disease in this group, and has been used as a model for comparison with other disorders as it affects the development and survival of sensory, sympathetic, and some parasympathetic neurons (2). Also called Riley-Day syndrome, FD was first described in 1949 based on five children who presented with defective lacrimation, excessive sweating, skin blotching, and hypertension (3). It is a devastating and debilitating disease that is present at birth with progressive neuronal degeneration throughout life. Since the original report, hundreds of papers have been published describing the various clinical manifestations as well as the marked variability of expression of the disorder. The following cardinal criteria have evolved for diagnosis of FD: absence of fungiform papillae on the tongue, absence of flare after injection of intradermal histamine, decreased or absent deep tendon reflexes, absence of overflow emotional tears, and Ashkenazi Jewish descent (1,2,4,5).

The loss of neuronal function in FD has many repercussions, with dysfunction in one system often leading to clinical problems in another (1,6-8). Gastrointestinal dysfunction is revealed shortly after birth by feeding difficulties such as poor suck, uncoordinated swallowing, and misdirection of bolus. Recurrent misdirection causes aspiration pneumonias and severe respiratory disease. FD patients also have abnormal respiratory responses to hypoxic and hypercarbic states and a high incidence of scoliosis that can restrict normal breathing. Gastroesophageal reflux is common and further increases the risk of aspiration in some patients. Approximately 40% of FD patients suffer severe vomiting crises that are associated with systemic autonomic dysfunction such as irritability, negative behavior, hypertension, tachycardia, blotchy erythema and excess sweating. Peripheral abnormalities of autonomic function include the distinctive lack of overflow tears with emotional crying, and corneal analgesia that places the patients at risk for corneal abrasions, ulceration and scarring. Abnormal vasomotor control is manifested by blotchy erythema during eating or emotional excitement, inappropriate sweating, and postural hypotension. Sensory deficiency is most notable as the inability to discriminate temperature differences and a dulled reaction to injury. The inability to differentiate between painful and light touch or contact is hampered and worsens with age. However, while bone and skin pain are poorly perceived, sensitivity for visceral pain is intact. Despite recent advances in the management of FD, the disorder is inevitably fatal with only 50% of patients reaching 30 years of age.

The clinical features of FD are all due to an autosomal recessive genetic defect that causes a striking progressive depletion of unmyelinated sensory and autonomic neurons (8-12). This neuronal deficiency begins during development as extensive pathology is evident even in the youngest subjects. Fetal development and postnatal maintenance of dorsal root ganglion (DRG) neurons are abnormal, significantly decreasing their numbers and resulting in DRG of grossly reduced size. Similarly, lateral root entry zones and Lissauer's tracts within the spinal cord are severely depleted of axons. Slow progressive degeneration is evident in continued neuronal depletion with age and in an increase in the number of abnormal residual nodules of Nageotte in the DRG. Loss of dorsal column myelinated axons is evident in older patients. The sural nerve is reduced in area and contains markedly diminished numbers of non-myelinated axons, as well as diminished numbers of small diameter myelinated axons. In the autonomic nervous system, superior cervical sympathetic ganglia are also reduced in size due to a severe decrease in neuronal population. The intermedio-lateral gray columns of the spinal cord also contain a low number of neurons and ultrastructural examination of peripheral blood vessels reveals absence of autonomic nerve terminals. The parasympathetic sphenopalatine ganglia are consistently reduced in size with low neuronal counts, but the other parasympathetic ganglia are minimally affected. These pathological findings have led to the suggestion that the FD gene, named *DYS*, may encode a member of the growing family of neurotrophic factors or their receptors which are crucial to the embryonic development and the postnatal survival of neurons (13-16). All of the family members identified to date can be excluded based on their chromosomal localization, but it is not inconceivable that the FD gene may be an as yet undiscovered neurotrophin or receptor, or a downstream participant in a related signal transduction pathway. Several investigators have also reported varied biochemical abnormalities in FD, but these are unlikely to represent the primary defect, probably reflecting instead secondary effects of the disorder (17-19).

Regardless of what the FD gene may be, this will likely be the first hereditary sensory neuropathy for which the genetic defect is discovered. Although there is considerable overlap, the diagnostic criteria for FD permit it to be distinguished from at least five other known congenital sensory neuropathies (2,4). The hereditary sensory neuropathies can be broadly subdivided into two classes. The three members of the first class, FD, congenital sensory neuropathy with anhydrosis (HSN IV), and progressive panneuropathy with hypotonia all show incomplete sensory loss with lack of axon flare following intradermal histamine, decreased corneal reflexes, lack of fungiform papillae on the tongue, and variable deep tendon reflexes. Factors often used to distinguish between these disorders are variability in muscle tone, intelligence, sensitivity to visceral pain, overflow tears and ethnic background. The second class of sensory neuropathies includes HSN II, HSN with skeletal dysplasia, and congenital autonomic dysfunction with universal pain loss, all of which show complete sensory loss. Despite the documentation of specific clinical criteria for each of the sensory neuropathies, their extreme variability of symptomatic expression can lead to diagnostic complications. Discovery of the genetic defect in FD will permit resolution of this disorder's relationship to the other sensory neuropathies, may well provide direct clues to the genetic causes of these similar disorders, and will almost certainly yield valuable insight into the processes involved in normal development and maintenance of the sensory nervous system.

FD is an autosomal recessive disease that is currently limited to the Ashkenazi Jewish population (5,20,21). However, as Ashkenazi descent has become one of the diagnostic criteria for the disorder, there may well be related cases among the sensory neuropathies in other populations, particularly in view of several reports of 'variant' FD in non-Jewish patients. Clearly, discovery of the FD defect will allow us to determine whether in fact these other diseases are allelic. FD is a very common disorder in the Ashkenazim with a carrier frequency of 1 in 30 (22). In 1990, we set out to determine the chromosomal localization of the FD gene (*DYS*) by genetic linkage analysis, with a mandate from the Dysautonomia Foundation to develop a prenatal diagnostic test for the disorder. In collaboration with Dr. Felicia Axelrod at the Dysautonomia Treatment and Evaluation Center at NYU Medical Center and Dr. Channa Maayan at Hadassah University Hospital in Israel, we were able to analyze 26 multiple-affected families. Initially, we genotyped these families with 191 genetic markers and effectively excluded 60% of the autosomal genome as the site of the defect (23). Analysis with the EXCLUDE program then predicted that *DYS* might be located on chromosomes 2, 4, 5q, 9, or 10. We concentrated on chromosome 9 as the best candidate from our initial studies and discovered the location of *DYS* in 1993 (24). The FD defect was completely linked to the marker *D9S58* with a lod score of 21. Genotyping of additional markers allowed us to confine the gene to an 11 cM region between *D9S53* and *D9S105* in 9q31-33. With the continued definition of closer flanking markers, we were able to supply the Dysautonomia Foundation with the requested prenatal linkage test for at risk families. This test has been successfully implemented with appropriate genetic counseling by groups at New York University and Mount Sinai Hospital in New York City and at the Hadassah Hospital in Jerusalem (25,26).

In view of the ethnic isolation of FD among the Ashkenazim, we have examined the closely linked markers for evidence of allelic association that would indicate descent from a genetic founder. The disequilibrium observed with the '18' allele at *D9S58* is striking; this allele is present on 73% of affected chromosomes and only 5% of control chromosomes(24). Since the initial discovery of linkage, we have worked to identify new genetic polymorphisms in the FD region to aid in the construction of a complete haplotype and to enable us to better define the candidate region. Recently we have localized the marker *D9S1677* to the FD region and have shown that the level of disequilibrium seen with this marker is much higher than with *D9S58*; 95% of FD chromosomes have a '12' (or related allele) at this marker compared with only 2.7% of control chromosomes. Using these markers, we have built a conserved haplotype that spans more than 1.1 Mb and suggests that >98% of all FD chromosomes carry the same genetic mutation. Of 286 FD chromosomes examined in detail, only 5 appear to be unrelated to the common haplotype, suggesting that they may possess different mutations. Until recently, the definitive *DYS* candidate region based on actual observed recombination events in families was a 3 cM interval between *D9S748* and *D9S105*. However, by examination of the haplotypes and inferred ancestral recombination events and by delineation of a crossover in a new FD family, we have been able to shrink this interval to <2 cM. We are currently working to isolate additional genetic markers that will enable

us to further narrow this interval. Our other efforts have been focused on the construction of a complete physical map that spans the entire *DYS* candidate region. We have constructed a YAC contig that spans 1.1 Mb and covers the region most likely to contain the FD gene. These YACs were used to identify cosmids that have been used to both detail the physical map and to identify candidate genes through the method of exon amplification (27-30). To date several candidate genes have been isolated and assessed (see Preliminary Results).

The work performed to date aimed at development of a prenatal diagnostic test for FD has yielded the information necessary to move forward with isolation of the disease gene by positional cloning. In the past few years, this strategy has grown in prominence in human genetics as it represents a straightforward approach to genetic disorders whose investigation has proved intractable by other strategies. Indeed, the positional cloning approach has identified genetic defects for a wide variety of important disorders, including cystic fibrosis, Huntington's disease, breast cancer, Alzheimer disease and many more (31). In each case, discovery of the disease gene has enabled researchers to move toward delineating the mechanism of pathogenesis, with the aim of building the foundation of knowledge necessary to define and test rationale treatments. Compared with many of these other disorders, FD represents a far better candidate for the positional cloning paradigm. It can be approached genetically as a very homogeneous disorder, thereby permitting the candidate region to be narrowly limited. Given the clear evidence for a common founder haplotype and our previous success at identifying both simple sequence repeat markers (SSR) and single strand conformational polymorphisms (SSCP), we are confident that we will be able to significantly reduce the *DYS* candidate region. Advances in cloning of large fragments of human DNA and in identification of human genes within cloned sequences make it very feasible to isolate the entire candidate region as overlapping clone sets and to identify and screen candidate genes. Thus, based on our experience in identifying the Huntington's disease gene (32), the neurofibromatosis 2 gene (33) and the Batten's disease gene (34) using similar approaches, we are also confident that the FD defect will be identified within the period of this grant.

The impact of isolation of the FD defect will be multi-fold. First, it will allow the prenatal diagnostic test to be extended beyond families with a previous FD child. Indeed, given the evidence for a shared mutation on almost all FD chromosomes, an effective carrier test that could be used if desired for population screening in the Ashkenazim is a strong likelihood. Second, the identification of the disease gene will provide the means for direct comparison of the other sensory neuropathies with FD, allowing a classification based on the primary genetic cause rather than subtle symptomatic differences. It will reveal whether other mutations at the same locus are a significant cause of sensory neuropathy in non-Jewish populations and may yield related genes that could be targeted for mutational screening in this family of disorders. Third, knowledge of the nature of the *DYS* gene, of its protein product, and of the effect of its likely functional deficiency in FD will set investigations of this disorder on the path toward an effective therapy for preventing the progressive neuronal loss. The disorder might prove amenable to protein replacement, like Gaucher disease, or more complex genetic or pharmacologic strategies may be necessary. In any event, the drive in modern biomedical research toward development of techniques for genetic therapy and rational drug design is a very strong one, being carried on in many different fields. The advances to be expected from such research will obviously benefit many different genetic disorders, but can only be applied to a disorder like FD once the primary defect has been ascertained.

Finally, although FD represents a tragic affliction that imposes suffering and early death on its victims, it also has the potential to provide a new and different window on a very fundamental biological problem, the development and maintenance of sensory neurons. This is a very active area of research that is currently being examined by a number of groups using targeted disruption of the genes for neurotrophins and their receptors in transgenic mice (35-37). FD represents an experiment of nature in which a similar mutation has been introduced in man. Discovery and characterization of the gene that has been disrupted will complement the knowledge being gained from the mouse system. The FD gene does not encode one of the mapped growth factors or receptors and, on the basis of the experience in other disorders that have succumbed to positional cloning, it may well prove to encode a novel protein. Identification of this protein will likely pinpoint a new participant in the web of biochemical interactions necessary for neuronal development and survival.

c. Preliminary Results

c.1 Publications resulting from our work on familial dysautonomia

Papers:

Blumenfeld A, Axelrod FB, Trofatter JA, Maayan C, Lucente DE, Slaugenhaupt SA, Liebert CB, Ozelius LJ, Haines JL, Breakefield XO, Gusella JF. 1993. Exclusion of familial dysautonomia from more than 60% of the genome. *J. Med. Genet.* 30:47-52.

Blumenfeld A, Slaugenhaupt SA, Axelrod FB, Lucente DE, Maayan C, Liebert CB, Ozelius LJ, Trofatter JA, Haines JL, Breakefield XO, Gusella JF. 1993b. Localization of the gene for familial dysautonomia on chromosome 9 and definition of DNA markers for genetic diagnosis. *Nature Genetics*. 4:160-164.

Breakefield XO, Ozelius L, Bothwell MA, Chao MV, Axelrod F, Kramer PL, Lanahan AA, Johnson DE, Ross AH Gusella JF. DNA polymorphism for the nerve growth factor receptor gene exclude its role in familial dysautonomia. *Mol Biol Med* 1986; 3:483-494.

Eng CM, Slaugenhaupt SA, Blumenfeld A, Axelrod FB, Gusella JF, and Desnick RJ. 1995. Prenatal diagnosis of familial dysautonomia by analysis of linked CA-repeat polymorphisms on chromosome 9q31-q33. *Am. J. Med. Genet.* 59: 349-355.

Oddoux C, Reich E, Axelrod F, Maayan C, Blumenfeld A, Slaugenhaupt SA, Gusella JF, Ostrer H. 1995. Prenatal diagnostic testing using for familial dysautonomia using linked genetic markers. *Prenatal Diagnosis* 15:817-826.

Ozelius L, Schuback DE, Stefansson K, Slaugenhaupt SA, Gusella JF, Breakefield XO. 1992. Dinucleotide repeat polymorphism for the hexabrachion gene (*HXB*) on chromosome 9q32-34. *Hum Mol Genet* 1:141.

Sahin M, Slaugenhaupt SA, Gusella JF, and Hockfield, S. 1995. A Protein Tyrosine Phosphatase, *PTPH1*, is expressed in a segment-specific pattern in the rat diencephalon. *Proc. Natl. Acad. Sci. USA*. 92:7859-7863.

Slaugenhaupt SA, Blumenfeld A, Liebert CB, Mull J, Lucente DE, Monahan M, Breakefield XO, Axelrod FB, Maayan C, Parada L, and Gusella, JF. 1995. The human gene for neurotrophic tyrosine kinase receptor type 2 is located on chromosome 9 but is not the familial dysautonomia gene. *Genomics*. 25:730-732.

Slaugenhaupt S, Liebert CB, Mull J, MacCormack K, LeBel A, Buckler A, McCormick MK, Ekong R, Gusella J. 1996. The gene for tumor necrosis factor receptor associated factor 2 maps to human chromosome 9q34. (in prep).

Slaugenhaupt S, Liebert CB, Mull J, MacCormack K, LeBel A, Buckler A, McCormick MK, Blumenfeld A, Gusella J. 1996. Evaluation of *PTPH1*, a protein tyrosine phosphatase, as a candidate gene for familial dysautonomia. (in prep).

Slaugenhaupt S, Liebert CB, Mull J, MacCormack K, LeBel A, Buckler A, McCormick MK, Ekong R, Gusella J. 1996. Localization and genomic structure of the human *PTPA* gene to chromosome 9q34. (in prep).

Abstracts:

Blumenfeld A, Slaugenhaupt SA, Lucente DE, Trofatter JA, Axelrod FB, Maayan C, Liebert CB, Ozelius LJ, Haines JL, Breakefield XO, Gusella JF. 1992. Exclusion of the familial dysautonomia gene from 75% of the genome. *Am. J. Hum. Genet.* 51suppl:A360.

Blumenfeld A, Slaugenhaupt SA, Lucente DE, Liebert CB, Axelrod FB, Maayan C, Monahan M, Ozelius LJ, Trofatter JA, Haines JL, Breakefield XO, and Gusella JF. 1993. The gene for familial dysautonomia is linked to chromosome 9 and shows strong linkage disequilibrium with *D9S58*. Second International Chromosome 9 Workshop, Chatham, MA.

Blumenfeld A, Slaugenhaupt SA, Lucente DE, Monahan M, Axelrod FB, Liebert CB, Maayan C., Haines JL, Breakefield XO, and Gusella JF. 1993. The familial dysautonomia gene maps to chromosome 9q31-33 and shows strong allelic association with *D9S58*. *Am. J. Hum. Genet.* 53suppl:A979.

Kwiatkowski DJ, Dib C, Slaugenhaupt SA, Povey S, Gusella JF, Haines JL. 1993. An index marker map of chromosome 9. Second International Chromosome 9 Workshop, Chatham, MA.

Kwiatkowski DJ, Dib C, Slaugenhaupt S, Povey S, Gusella JF, and Haines JL. 1993. An index marker map of chromosome 9 provides strong evidence for positive interference. *Am. J. Hum. Genet.* 53suppl:A259.

Liebert CB, Slaugenhaupt SA, Blumenfeld A, Lucente DE, Gusella JF. 1993. Identification of Simple Sequence Repeat (SSR) Polymorphisms on Human Chromosome 9. Second International Chromosome 9 Workshop, Chatham, MA.

Lucente DE, Blumenfeld A, Slaugenhaupt SA, Liebert CB, Monahan M, Trofatter JA, Haines JL, Gusella JF. 1993. A high resolution map of 9q22.3-q33. Second International Chromosome 9 Workshop, Chatham, MA.

Slaugenhaupt SA, Blumenfeld A, Arnold J, Liebert CB, Mull J, Lucente DE, Monahan M, McCormick MK, Gusella JF. 1993. Physical Mapping of the familial dysautonomia gene region of human chromosome 9. Second International Chromosome 9 Workshop, Chatham, MA.

Slaugenhaupt SA, Blumenfeld A, Liebert CB, Monahan M, Mull J, LeBel A, MacCormack K, Breakefield XO, Axelrod F, Maayan C, and Gusella JF. 1994. Isolation of candidate genes and physical mapping in the familial dysautonomia region of chromosome 9q31. *Am. J. Hum. Genet.* 55suppl:A31.

Slaugenhaupt SA, Blumenfeld A, Liebert CB, Mull J, Lucente DE, Monahan M, Buckler AJ, McCormick MK, Breakefield XO, and Gusella JF. 1993. Physical mapping of the familial dysautonomia gene region of human chromosome 9. *Am. J. Hum. Genet.* 53suppl:A1360.

Slaugenhaupt SA, Blumenfeld A, Liebert CB, Mull JM, MacCormack K, LeBel A, O'Leary K, Lucente D, McCormick MK, Buckler A, Gusella JF. 1995. A Physical Map Surrounding the Familial Dysautonomia Gene. Fourth International Workshop on Chromosome 9, Williamsburg, VA.

Slaugenhaupt SA, Liebert CB, Mull J, MacCormack K, LeBel A, O'Leary K, Lucente D, McCormick MK, Buckler AJ, Maayan C, Axelrod F, Blumenfeld A, Gusella JF. 1995. Evaluation of Candidate Genes for Familial Dysautonomia. *Am. J. Hum. Genet.* 57suppl: A270.

Slaugenhaupt SA, Anderson MA, Jenkins B, MacCormack K, Lerner T, Gusella JF. 1992. Generation of a radiation hybrid mapping panel for chromosome 9. *Chromosome 9 Workshop*, Cambridge, England.

c.2 Introduction

The MGH Molecular Neurogenetics Unit, under Dr. Gusella's direction, has a long history of genetic linkage analysis and positional cloning in human neurologic disease. In addition to finding the first linked DNA marker for a previously unmapped autosomal disorder, that for Huntington's disease (38) the group has participated in the initial genetic mapping of a variety of other diseases, including neurofibromatosis 1 (39), neurofibromatosis 2 (40), von Hippel Lindau disease (41), familial Alzheimer disease (42), dystonia musculorum deformans (43), hyperkalemic periodic paralysis (44) and achondroplasia (45). The Unit has also developed the genomic expertise and resources to identify genetic defects based on their chromosomal location, including the genes causing Huntington's disease (32), neurofibromatosis 2 (33), Batten's disease (34), Wilson's disease (46) and paramyotonia congenita (47). The group has had a long interest in familial dysautonomia and took the approach of excluding candidate genes, such as the nerve growth factor and its receptor (13,16) until sufficient families and human linkage markers became available to attempt a scan of the entire genome for the site of the defect. In 1991, the Dysautonomia Foundation awarded a contract to develop a prenatal genetic test for dysautonomia using the linkage strategy. The current proposal is the outgrowth of the work done to develop the prenatal linkage test. We successfully identified the location of the FD gene and have provided close flanking markers for an accurate prenatal test applicable to at risk families. The effort to positionally clone the FD gene has proceeded well, but it is straining the resources of the Dysautonomia Foundation and goes far beyond the mandate supported by the contract, which ends in 1996. Consequently, this section, Preliminary Results, will detail the work that we have done leading up to the current proposal which has placed us at the brink of identifying the FD defect and which we hope to bring to fruition with NIH support. This would represent the first gene defect isolated among the sensory neuropathies and will be of value both for gaining an understanding of this important group of neurogenetic disorders and for studying the fundamental biology of pain and sensation. Throughout this Preliminary Results section, references to publications of our work that forms the foundation of this proposal are referred to by author/journal and can be found in the above list of papers and abstracts. Other references are included as usual in section g. Literature Cited

c.3 FD patient samples

The successful application of a positional cloning strategy depends largely on the quality and amount of family material available for study. FD is a disorder affecting the Ashkenazi Jewish population, with most cases therefore being seen in either the U.S. or Israel. Our work has been carried out with samples from two major sources. Dr. Felicia Axelrod directs the Dysautonomia Diagnostic and Treatment Center at the New York University Medical Center, and is the recognized world expert on diagnosis and management of this disorder. Dr. Channa Maayan who has a similarly long experience in dealing with FD sees Israeli FD families at the Hadassah University Hospital in Jerusalem, Israel. She is aided in this work by Dr. Anat Blumenfeld, a former post-doctoral fellow in the Gusella laboratory, who has continued to collaborate on the project since her return to Israel. Together, these physicians have access to the vast majority of diagnosed FD cases and they have supplied us with blood samples for our studies.

The blood is processed to prepare direct DNA spotted on filters as a back-up for later extraction of limited DNA for PCR, and transformed with EBV to generate lymphoblast cell lines as permanent sources of DNA, RNA and protein for analysis. To date, we have initiated 496 cell lines from FD patients and their family members. Among these samples are 26 multicase families that were used to establish the chromosomal location of the FD defect (Blumenfeld et al., J Med Genet 1993; Blumenfeld et al., Nature Genetics 1993). The remaining samples of single affected individuals are used for haplotype analysis to narrow the location of the defect. We continue to receive additional samples as new patients are ascertained, particularly from Israel due to the influx of Russian émigrés. These collaborators also supply samples from 'variant' FD and other sensory neuropathies for analysis after the FD defect is found.

c.4 Exclusion studies and discovery of linkage

Our initial effort to identify a linked marker for FD used 23 of the multicase families then available (Blumenfeld et al., J Med Genet 1993). Genotyping was performed with 191 restriction fragment length polymorphism (RFLP) and simple sequence repeat (SSR) markers spread throughout the genome. No significant lod score was identified, but we estimated that we had excluded the FD gene from 60% of the autosomal regions. Analysis with the EXCLUDE program predicted segments of chromosomes 2, 4, 5q, 9 and 10 as the most promising remaining possibilities.

In following up on these analyses, we focused first on 9q where only a portion of the arm remained to be excluded. We tested the *HXB* (hexabrachion) gene from 9q32-q33 (Ozelius et al., Hum Mol Genet 1992) and immediately achieved a significant positive lod score of 9 at a recombination fraction of 0.04 . Follow-up studies of additional markers yielded significant positive lod scores with 10 loci. The maximum lod score achieved was 21.1 with no recombinants at *D9S58*. Multi-point analysis indicated the disease gene was located between flanking markers *D9S53* and *D9S105*, in an 11 cM interval in 9q31 containing *D9S58* (Blumenfeld et al., Nature Genetics 1993).

c.5 Delineation of Founder Effect

Many disorders that show a markedly higher gene frequency in a particular race or ethnic group may do so because of a founder effect. In this case most disease genes present in the population carry the same ancestral mutation. The linkage data strongly supported this view in FD. At *D9S58*, 73% of disease chromosomes carried an '18' allele although this allele is found on only 5 % of control chromosomes. Three distal markers, *D9S59*, *D9S105* and *D9S106*, mapping 0.5-1.5 cM from the defect also displayed significant linkage disequilibrium with FD. This clear founder effect, detectable over more than 1 cM, suggests that the ancestral FD mutation may be only several hundred years old (21).

c.6 Prenatal Diagnosis

Under the auspices of the Dysautonomia Foundation, a prenatal linkage test for FD using the flanking markers that we have identified has been implemented at the NYU Medical Center and at the Mount Sinai Hospital in New York City and at the Hadassah University Hospital in Jerusalem (Oddoux et al., Prenatal Diag 1995; Eng et al., Am J Hum Genet 1995). These markers are sufficient to yield an informative predictive result in most families at 25% risk with the proband available for testing. Indeed, the allele association observed in FD has made some prediction possible even in families for which no proband DNA can be obtained. To date, 46 fetal predictions have been carried out successfully among the three centers. At the New York hospitals, 30 predictive test have been performed leading to 18 healthy newborns, 9 terminations due to a positive test and 3 miscarriages. Moreover, the strong founder effect observed with the initial linkage markers and closer markers isolated subsequently (see below) suggests that when the FD gene is found, risk assessment will be straightforward even for Ashkenazi Jewish individuals with no known family history of FD.

c.7 Candidate genes

Once the location of the FD gene had been ascertained, it was possible to consider the possibility that the defect might lie in a gene already mapped to chromosome 9q. The most obvious candidates were genes encoding neurotrophins, their receptors, or any proteins that might participate in a neurotrophin-related signal transduction pathway. Most genes in the first two categories had already been mapped to other chromosomes, excluding them from causing FD. For the remainder, we probed DNA from two human-hamster somatic cell hybrid lines, GM10611 and 640-63a12, containing only human chromosome 9 and 9q, respectively for evidence of a related locus on this chromosome (48,49). Of several probes used, one, a mouse trkB probe gave a positive signal for both hybrids. This suggested that the unassigned homologous human locus, *NTRK2*, mapped to chromosome 9. This gene encodes a neurotrophin receptor whose preferred ligands are brain-derived neurotrophic factor (BDNF) and neurotrophin-4 (NT-4). We screened a gridded chromosome 9 cosmid library with the mouse probe and used exon amplification to isolate coding sequences from resulting cosmid clones. Four different exonic products were obtained with strong homology to both mouse and rat trkB genes, confirming that the human

NTRK2 candidate gene did indeed map on chromosome 9. To test for linkage with *DYS*, we screened for and identified a dinucleotide repeat locus common to three of the overlapping cosmids. This SSR proved to be polymorphic and informative for linkage in both reference and FD families. To our dismay, this strong candidate gene mapped in 9q21-q22, more than 20 cM proximal to *DYS* (Slaugenhaupt et al., Genomics 1995).

Two other candidate genes emerged from a chromosome 9 genome mapping project being carried out by Alan Buckler in the Molecular Neurogenetics Unit. Dr. Buckler had generated a library of random exon products from chromosome 9 (29). In sequencing these products, he identified exons from two unassigned genes not known to map to chromosome 9. One encodes *PTPA*, a protein that stimulates the tyrosyl phosphatase activity of protein phosphatase 2A. This protein plays a crucial role in the regulation of many cellular processes including growth signaling, a function not incompatible with the possible role of *DYS*. The other encodes *TRAF2*, or tumor necrosis factor receptor associated factor II, a protein associated with the cytoplasmic domain of the tumor necrosis factor receptor type 2 (TNF-R2) that is required for signal transduction. *TRAF2* could potentially also associate with other receptors and might function in a subsequent step in a signaling cascade whose interruption can result in cell death, making it a candidate for *DYS*. Both of these genes can now be excluded based on physical mapping within 9q34. Manuscripts describing the mapping and analysis of these two genes are currently in preparation (Slaugenhaupt et al.).

c.8 Physical Mapping

With the knowledge that *DYS* is between *D9S53* and *D9S105* and shows very strong linkage disequilibrium with *D9S58*, we set out to generate a physical map of the candidate region. The strategy was to expand the amount of cloned DNA around the markers and to use this material to generate new polymorphic markers. These new polymorphisms were then used to narrow the minimum candidate region by observation of actual crossovers in families and by inference of historical crossovers by haplotype analysis. We began by isolating YAC clones containing the markers *D9S58* and *D9S105* from complete digest non-chimeric YAC libraries prepared from flow-sorted chromosome 9 DNA by Dr. Mary Kay McCormick (50). We then initiated a bi-directional walk from the *D9S58* YACs using an Alu-PCR walking strategy (51-54). Specifically, Alu-PCR products generated by amplifying DNA located between Alu interspersed repeat elements were generated from the *D9S58* YACs and used to probe gridded arrays of similar Alu-PCR products from both the CEPH YAC library and chromosome 9 YAC libraries (50,55). The YACs were then used to isolate cosmids from arrayed chromosome 9-specific cosmid libraries that could be used to develop a minimal overlapping sets of cosmids (56-58). Additional walking steps were performed in the YAC libraries using new Alu-PCR products from the YACs or the corresponding cosmids (51-54). Occasionally, walking was also performed by screening the cosmid or YAC libraries with sequences recovered from YAC or cosmid ends (59).

As the cosmids were identified, they were processed by the exon amplification technique to generate exonic probes (27-30). While the exon trapping procedure was intended to facilitate finding genes in a candidate region, it also provides a very efficient way to map the region (60,61). The exon clones are almost invariably small and single copy, making this a very rapid method for generating single copy hybridization probes from cloned cosmids without the need to screen out repeat sequences. Similarly, it is straightforward to convert these single copy stretches to reliable PCR assays for STS mapping of YACs and cosmids. Consequently, the physical map was checked, and cosmid overlaps detected by matching all cloned exon products against the cosmids corresponding to the YACs in the region.

This physical mapping is still continuing, but the product to date is shown in Figure 1. This diagram depicts a 1.1 Mb YAC contig consisting of only those YAC clones most useful for building the map. Also shown is a minimal cosmid contig (selected from >300 cosmids) that spans more than 700 kb surrounding *D9S58*. All of the mapping reagents, including polymorphisms, STSs, YAC and cosmid end-clones and exons, were used to orient the map and to develop overlaps. To date 75 independent exon clones have been isolated across this region and sequenced for comparison with GenBank and dbest databases.

Figure 1- Current physical map of the familial dysautonomia gene region of human chromosome 9q31. This schematic diagram shows a 1.1 Mb YAC contig (bold lines) and a minimal cosmid contig (fine numbered lines) that spans >700 kb. The scale of the map is in kb and the centromeric and telomeric directions are given on the left hand side of the map. Shown above the contig are polymorphisms (**), STSs (names beginning with 's'), YAC and cosmid end clones ('Ye' and 'Ce' names, respectively) and exons (names beginning with 'x').

c.9 *PTPH1* as a candidate gene

Nine exons obtained from our initial exon amplification experiments around *D9S58* revealed sequence identity in database searches to a published gene, *PTPH1*, which was a viable candidate for *DYS*. In addition, isolation of 2 cDNA clones in our laboratory demonstrated that *D9S58*, the marker displaying the highest linkage disequilibrium with FD, is located in the 3' UTR of *PTPH1*. *PTPH1* encodes a protein tyrosine phosphatase that contains a protein 4.1 superfamily domain, suggesting that it may be targeted to an integral membrane protein (62). *PTPH1* could be a good candidate for *DYS* due to the apparent involvement of phosphatases in cell development and survival. Indeed, *in situ* hybridization experiments performed in collaboration with Dr. Susan Hockfield provided potential support for the *PTPH1* candidacy. Unlike other protein tyrosine phosphatases, *PTPH1* showed relatively restricted expression, with prominent signal in the developing dorsal root ganglia and in the sensory nuclei of the adult thalamus, regions expected to be involved in the manifestation of the FD phenotype (Sahin et al., PNAS 1995).

The published cDNA is 3984 bp, with 2739 bp of open reading frame and a 3' untranslated region of 1222 bp (62). The open reading frame predicts a protein of 913 amino acids that can be described in terms of three segments. 1) The N-terminal segment displays homology to band 4.1, ezrin, and talin. 2) The central segment bears putative phosphorylation sites for protein-serine/threonine kinases. 3) The C-terminal segment is homologous to members of the protein-tyrosine phosphatase family. Given the strong disequilibrium with *D9S58* in the 3'UTR and the promising *in situ* hybridization results, we performed a detailed analysis of *PTPH1* in FD patients and controls.

Using both SSCP and direct sequencing of RT-PCR products, we found no evidence for a pathogenic mutation in the known *PTPH1* coding sequence. However, we did identify 3 polymorphisms, one of which alters an amino acid. Incorporation of these polymorphisms into our haplotype analysis and the absence of mutations targeted our efforts to the 5' portion of *PTPH1*. We determined that the gene consists of at least 25 exons spanning 80 kb. However, we isolated a trapped exon containing the presumed initiator ATG. This indicated that this sequence was an internal exon since exon 1 of a gene cannot be trapped by the exon amplification technique. We utilized 5' RACE PCR and successfully cloned two different 5' cDNAs (63). Both products map upstream in the cosmid contig, contain very GC rich regions and splice at base pair 7 of the published sequence. At least one of the alternative 5' ends would predict additional amino acids at the N-terminus of the protein.

The failure to identify a pathogenic mutation in the coding sequence was disappointing, but did not exclude *PTPH1* as a candidate gene as a mutation outside the known coding sequence might be involved. We turned to analysis of the protein to determine whether *PTPH1* is in fact expressed in FD patients. The *PTPH1* protein was examined by Western blot analysis in collaboration with Dr. Nicholas Tonks of the Cold Spring Harbor Laboratory who has generated specific polyclonal and monoclonal antibodies against the protein. Our first attempts at detecting *PTPH1* protein in lysates from lymphoblast cell lines with *PTPH1* antibodies were unsuccessful due to the low expression of the gene. However, we obtained valuable autopsy material from an FD patient through collaboration with Dr. Felicia Axelrod, Dr. Bruce Quinn, and Dr. John Pearson at NYU that enabled us to perform the critical experiment. Indeed the *PTPH1* protein was expressed in FD thalamic and adrenal tissues in a pattern comparable to control tissue. This result, combined with our negative results in mutation scanning, reduced our enthusiasm for *PTPH1* as a *DYS* candidate (Slaugenhaupt et al., in preparation).

Figure 2-Genetic map of the *DYS* region. An ideogram of human chromosome 9 is shown with the genetic map of the DYS region in 9q31. The 3 cM interval between *D9S748* and *D9S105* is expanded on the right to show all polymorphic markers and the location of the minimal FD candidate region.

c.10 Genetic mapping and haplotype analysis

Throughout our generation of the physical map surrounding *D9S58*, we have continued to generate SSR markers from the cosmid clones obtained and to assess any SSR markers reported by others for 9q31-q32. Our aim has been to push the genetic analysis to its limit, to precisely define the minimum candidate region and thereby minimize the number of genes that must be analyzed in detail. The candidate region originally defined by our 26 multicase families spanned 11 cM between *D9S53* and *D9S105*. Better mapping of these markers in recent years has indicated that 7 cM is a better estimate for this interval. However, the candidate region was narrowed to 3 cM by the addition of *D9S748* (64)which proved to be a closer flanking marker on the centromeric side (see Figure 2). We reasoned that it was unlikely that additional recombinants could be defined as few additional multicase families could be expected. However, the linkage disequilibrium noted for several markers offers a supplemental strategy for narrowing the candidate region since it reflects inheritance of a founder mutation. Examination of decay in the founder haplotype surrounding the disease gene can then be used to infer the location of historical recombination events and thereby narrow the location of the defect.

In our initial analysis with *D9S58* and *D9S105*, it was clear that the haplotype '18-8' was overrepresented on FD chromosomes (Blumenfeld et al, Nature Genetics 1993). The proximal flanking marker, *D9S748*, shows only weak allele association with *DYS*, suggesting multiple crossovers on the centromeric side. To saturate the FD region with genetic markers, we scanned all public databases for SSR polymorphisms while at the same time screening our expanding physical map for di-, tri- and tetranucleotide repeats. A more complex haplotype was constructed using our new markers, *D9S309*, *D9S310* and *D9S311*, along with *D9S160* (65). None of these loci shows recombination with *DYS* in any of our multicase families, but the haplotypes from a set of 180 FD chromosomes, shown in Table 1, exhibit clear evidence of historical crossovers. The FD chromosomes subdivide into three conserved haplotypes for *D9S310-D9S309-D9S58*: '5-1-18', '6-2-13' and '6-2-16'. Together, these three haplotypes account for 91% of this set of FD chromosomes. The remaining 9% exhibit different haplotypes seemingly unrelated to these three. The existence of several different pairs of alleles at *D9S160* and

D9S311 in association with the major '5-1-18' haplotype indicates that crossovers have occurred between *D9S58* and *D9S160*. These data strongly suggest that *D9S160* is a distal flanking marker for *DYS*.

Table 1- Core marker haplotypes on FD and non-FD chromosomes

D9S310	*D9S309*	*D9S58*	*D9S160*	*D9S311*	% FD chromosomes	% Control (non-FD) chromosomes
5	1	18	7	7	73	1
5	1	18	6	4	3	0
5	1	18	5	7	1	0
5	1	18	6	7	1	0
5	1	18	7	4	1	0
6	2	13	7	7	7	2
6	2	16	7	2	5	1
X	X	X	X	X	9	96

X indicates no match to the consensus haplotypes

Further interpretation of the relationship between the haplotypes requires consideration of four alternative possibilities:

1. Each of the three major haplotypes could represent an unrelated mutation from an independent founder.
2. The '5-1-18' haplotype could mark one founder chromosome while the '6-2-13' and 6-2-16' chromosomes could be ancestrally related to each other and represent a second independent mutation. The historical crossover relating '6-2-13' and '6-2-16' would place *DYS* proximal to *D9S58*.
3. The '5-1-18' haplotype could be ancestrally related to the '6-2-13' chromosome since both have '7-7' for *D9S160-D9S311*. This possibility is supported by the fact that the majority of the '5-1-18' chromosomes and 100% of the '6-2-13' FD chromosomes have an '8' allele at the more distant *D9S105*. This scenario would make *D9S58* a proximal flanking marker and confine *DYS* to the *D9S58-D9S160* interval.
4. All three haplotype could derive from a common founder but differ because of undetected historical crossovers proximal to *D9S310*. In this case, *DYS* would be located centromeric to *D9S310*.

It was particularly crucial to resolve these various possibilities since they each led to different predictions for the location of *DYS* proximal to *D9S160* and would therefore support different priorities for physical mapping and candidate gene assessment. Scenario 3 seemed most likely because the *D9S160-D9S311-D9S105* similarities between the '5-1-18' and '6-2-13' chromosomes represent a pattern not common on control chromosomes (4%). This hypothesis would also confine *DYS* to the small interval containing the *PTPH1* gene. However, although we had an excellent candidate in this interval, we recognized the need to formally resolve the 4 possibilities outlined above.

We continued our walk from *D9S58*, concentrating now on the proximal direction, in order to add new markers to the haplotype. The location of *DYS* was resolved by addition of several loci proximal to *D9S310*. In particular, a newly published dinucleotide repeat marker, *D9S1677* (66), which maps 250 kb proximal to *D9S58*, shows the strongest linkage disequilibrium with FD yet observed, strongly supporting scenario 4. Eighty-one percent of FD chromosomes have a '12' allele compared with only 2% of control chromosomes. An additional 14% of FD chromosomes have a '13' or '14' allele at this site. We believe that these alleles were generated by slippage of the '12' allele as they are present on <1% of control chromosomes. When *D9S1677* was incorporated into the FD haplotypes, it was clear that all three, '5-1-18', '6-2-13' and '6-2-16' are derived from a common founder with a '12' allele at *D9S1677*. In addition, a *D9S1677* '12' allele was also present on 17 of the 22 chromosomes that had previously appeared unrelated, suggesting that these too were derived from the same founder.

To supplement this haplotype analysis, we recently developed a pentanucleotide SSR marker, 131A7, that maps 50 kb proximal to *D9S1677*. All chromosomes with a '12' allele at *D9S1677* have the same allele for 131A7, as do the FD chromosomes with *D9S1677* '13' and '14' alleles. This confirms the conclusion that the '13' and '14' alleles are due to slippage from a '12' allele. In addition, these two markers indicate that 98% of the FD chromosomes have the same ancestral mutation. Only 5 chromosomes (representing only 3 novel haplotypes) might have independent mutations. As each of these haplotypes occurred paired with a classic founder haplotype in an affected individual, there is no concern of misdiagnosis. Only 2 control chromosomes match any of the three most frequent *D9S1677-D9S310-D9S309-D9S58-D9S160-D9S311*' haplotypes but both of these differ at *D9S105*. It should be noted that the control chromosomes for these analyses are normal chromosomes from FD carriers, so the usual carrier frequency of 1 in 30 is not a concern.

On the basis of the haplotype analysis alone, the distal limit of the *DYS* candidate region can be moved to *D9S310*. Recently, our collaborator Dr. Blumenfeld identified a recombinant in a new multicase Israeli family (which we have reconfirmed) that supports and extends this haplotype analysis. This family exhibits a recombination placing *DYS* proximal to *D9S58*. *D9S310* and *D9S309* were not informative in this family which was clearly non-recombinant for *D9S1677*. However, we have generated several less informative markers proximal to *D9S310* that were not included in the above discussion of haplotypes because of their low information content. One of these, 99E10, located 30 kb proximal to *D9S310* detects the new crossover and now represents the distal limit of the FD candidate region.

Thus, the FD defect maps in a segment of 9q31 bordered on the distal side by 99E10 and containing *D9S1677* (Figure 2). The proximal border has not yet been defined on the physical map. In formal terms, it is still at *D9S748*, 2 cM away. However, we expect that the conserved haplotype denoting the founder chromosome will decay on the proximal side of the disease gene at a rate comparable to the decay observed on the distal side. This would suggest that historical crossovers will be detected within the 300 kb proximal to *D9S1677*. Our major goal at present is to define new markers at the proximal end of the YAC contig to determine whether these set the proximal limit of the candidate region or whether another YAC walking step will be required.

c.11 Molecular Resources

We have accumulated a number of molecular resources that facilitate the pursuit of this project. These have come from outside investigators, and particularly from other faculty in the Molecular Neurogenetics Unit who have been involved in mapping of human chromosome 9. For rapid assignment of DNA probes to chromosome 9q, we have two reliable human-hamster somatic cell hybrid lines, containing only human 9 and 9q, respectively (48,49).

For YAC screening, we have the CEPH YAC library (55) and four different chromosome 9 specific YAC libraries (50). The latter were constructed from flow-sorted chromosome 9 DNA by Dr. Mary Kay McCormick at the Los Alamos National Laboratory and brought to the Molecular Neurogenetics Unit when she moved here as faculty. These libraries were constructed with ClaI, SacII, SalI and BamHI and have a far lower rate of chimerism than the CEPH library. The YAC libraries can be screened with individual PCR assays, but contig formation is far more efficient using Alu-PCR probes (51-54). Consequently, the Alu-PCR products from all the YAC libraries have been spotted on filters using a complex multi-dimensional pooling scheme that permits identification of individual YACs from the pattern of hybridization to a small number of filters.

To prepare cosmid contigs, we use two chromosome 9-specific libraries prepared using DNA from flow-sorted human chromosome 9 derived from a human-hamster hybrid containing this chromosome (56-58). One, provided by the Los Alamos National Laboratory, represents 5X coverage in superCOS and is stored in gridded 96 well arrays. It is screened by hybridization to 27 filters representing 9 well plates per filter. The second library was similarly prepared at the Lawrence Livermore National Laboratory as 5X coverage in LAWRIST16. It is screened with 20 filters with 15 plates/filter. We have also prepared an FD-specific superCOS library from a patient homozygous for the major founder haplotype.

To prepare trapped exons from the cosmid clones, we use the procedure developed by Dr. Alan Buckler, until recently a faculty member of the Molecular Neurogenetics Unit. Dr. Buckler was for the past two years engaged in generating chromosome-specific exon libraries by subjecting pools of anonymous cosmids to his technique (27-29). Consequently, the chromosome 9-specific pools of exons are also available for hybridization screening should these be needed. For cDNA isolation, we have accumulated a wide variety of cDNA libraries for use in our various positional cloning projects over the past few years. In our experience, almost all exon clones will yield a cDNA from screening a combination of fetal and adult human brain libraries. We also have fractions of a human brain cDNA library pooled according to a multi-dimensional scheme to permit rapid screening by PCR (67). Recently, we have also obtained the IMAGE consortium cDNA filters for more rapid screening based on gridded arrays (68).

c.12 Summary

Our work on familial dysautonomia has advanced over the years from a background project where limited progress was possible due to lack of families and genetic markers, through a targeted effort to develop an accurate prenatal diagnostic test to its current state, at the threshold of identifying the genetic defect. The discovery of the FD defect will have an impact on our understanding of this disorder, and sensory neuropathies in general. It will provide neuroscientists with new insights into development and maintenance of the sensory and autonomic nervous systems. Most importantly, however, it will provide hope to FD families by initiating the path toward a rational treatment for preventing progressive neuronal loss while providing the diagnostic technology to minimize transmission of the disease gene.

d. Research Design and Methods

d.1 Introduction

We believe that our preliminary work has provided a firm foundation for pursuing the identification of the *DYS* gene and that we have the clinical material, molecular biological reagents and experience in positional cloning to complete this task well within the term of this grant. The steps outlined in the following sections are relatively straightforward as they have proved successful in many similar projects both in this laboratory and elsewhere. The consequences of this phase of the work will be an answer to the question of what gene defect causes FD, a probable resolution to whether a growth factor signaling pathway is directly involved, and an assessment of whether similar defects are operative in other sensory neuropathies. It will also provide a carrier test applicable to the general Ashkenazi Jewish population in which the risk of unknowingly transmitting this recessive defect is extremely high.

The steps to be taken after the cloning of *DYS* are not as easy to predict, as they depend in large part on the precise nature of its protein product. However, in this research plan we will outline a variety of steps that can be taken to address the next obvious question, "How does the genetic defect cause the characteristic pattern of sensory neuropathy in FD?". The exact research plan for this phase of work will be decided upon when the *DYS* gene is cloned. The decisions will be guided by whether the protein product is completely novel or is related to a protein of known biological activity for which specific experimental systems have already been developed. In either event, the results of the second phase of this work will be an understanding of the expression of the *DYS* gene and identification or generation of model systems in which to study its physiological role. It is to be hoped that these advances will eventually lead to an effective rational treatment for FD, but it is certain that they will yield fundamental insights into both the normal and abnormal development and operation of the sensory nervous system.

d.2 Family samples

The FD DNA and cell lines available to us from our clinical collaborators currently provide more than 375 disease chromosomes, a number that we believe is sufficient to successfully isolate the disease gene. However, we will continue to solicit additional samples for two major reasons. First, identification of additional actual or historical recombination events could further narrow the minimum candidate region that must be scrutinized in detail. Second, examining more chromosomes might reveal additional haplotypes that could represent mutations different from that on the common founder haplotype. Each such chromosome would provide a comparison in which an independent alteration would be expected in the same gene. It would thus help in definitive identification of *DYS* and would provide the basis for genotype:phenotype correlations that might be revealing concerning protein function. In this regard, we will also continue to solicit patients with 'atypical' FD and with other sensory neuropathies for comparison with the classic FD defect. Our clinical collaborators will continue to collect autopsy tissues from FD patients whenever possible for biochemical and immunohistochemical analyses.

As each blood sample is received in the Molecular Neurogenetics Unit, it is processed in our Tissue Culture Core facility. DNA is extracted directly from blood using the Scotlab Nucleon II DNA extraction kit. For backup storage, drops of blood are spotted and dried on Schleicher & Schuell blood collection cards. The DNA in such spots survives for long periods of time with simple storage at room temperature and can be extracted as desired in quality and quantities sufficient for >25 PCR reactions (69). Finally, an aliquot of white cells isolated in lymphoprep tubes are transformed with EBV to establish a permanent cell line that can be used as a source of DNA, RNA or protein for detailed investigations (70). Each new DNA sample received is genotyped for core and flanking markers in the founder haplotype (*D9S748*, *D9S1677*, *D9S310*, *D9S58*). If necessary, the phase of alleles with respect to each other is deciphered by comparison with parents and/or siblings. In most cases, since the individual is homozygous for most or all of the founder haplotype, such comparisons are not needed. If a chromosome proves interesting because of variation from the consensus haplotype, more markers from the region are typed to better define haplotype relationships.

d.3 Building the physical map

The up-to-date status of the physical map of the immediate FD candidate region is shown in Figure 1. In addition to the YAC clones shown, we have isolated YACs corresponding to the flanking markers *D9S748* and *D9S105*, but these are not depicted as they do not overlap the contig and are at an uncertain physical distance. Work on the telomeric side of the 1.1 Mb contig has been discontinued as the new genetic evidence indicates that *DYS* is located on the centromeric side of the map. Currently, the cloned segment extends 320 kb centromeric to the distal border of the candidate region at 99E10. *D9S1677* lies within this candidate region, as does a less informative pentanucleotide SSR, 131A7, located 50 kb proximal to *D9S1677*.

As outlined in our preliminary results, the YAC contig is built by an Alu-PCR walking strategy (51-54), supplemented by YAC end-cloning (59), STS development and PCR screening of the YAC libraries as necessary. YACs are converted to cosmids by direct hybridization to cosmid arrays with suppression of repeat sequences (71). Cosmid contigs are constructed by hybridization with numerous exonic probes isolated by exon amplification. Where cosmid walking steps are needed, these are performed by the same Alu-PCR strategy as used with the YACs or by development of end-probes for hybridization screening. The entire map is verified and oriented by hybridization of exons, and PCR typing of STSs and SSRs. We are currently engaged in cosmid walking to close two gaps in the cosmid contig centromeric to 131A7, estimated at <100 kb and <30 kb, respectively.

We intend to continue defining overlapping clones centromeric to the end of the YAC contig at end-clone Ye13E8spc (Figure 1) only until a firm proximal border of the candidate region is defined by historical or actual crossovers. The lack of disequilibrium at the proximal flanking marker *D9S748* suggests that it will not be necessary to extend the map to this marker to identify such recombinations, but this will be done if necessary. In parallel with our YAC walking, we are converting the region of the contig centromeric to *D9S310* into BAC clones. This vector system provides us with an important alternative to YAC walking in the rare event that a YAC walking step should prove problematic. More importantly we will supply the BACs to Dr. Eric Lander and the MIT Genome Center who will perform large-scale sequencing of chromosome 9. Thus, we eventually expect to have the complete nucleotide sequence of the FD region.

d.4 Defining the genetic candidate region

The formal genetic candidate region defined by crossovers in multicase families extends from *D9S748* to 99E10, a genetic distance of <2 cM (Figure 2). There is every reason to believe that this interval can be narrowed further based on both actual recombinations in FD families and by historical crossovers defined by haplotype analysis. The discovery that 98% of FD chromosomes share the same or related alleles at *D9S1677* and 131A7 but differ for more telomeric markers reveals historical crossovers that occurred since the original FD mutation event. The overwhelming prevalence of the '12-5-1-18-7-7' haplotype for *D9S1677-D9S310-D9S309-D9S58-D9S160-D9S311* on FD chromosomes makes it most likely to be the founder haplotype on which the original ancestral mutation occurred. The other haplotypes were apparently generated by crossovers distal to *D9S1677* within this major founder haplotype.

Telomeric to *DYS*, we have a single actual crossover in a multicase family that occurred between *D9S1677* and 99E10. However, it is clear that several historical crossovers have also occurred in this interval based on the 26% of FD chromosomes that share the rare '12' (or related alleles) at *D9S1677* but differ from the '5-1-18' founder haplotype at *D9S310-D9S309-D9S58*. Among the frequently seen haplotypes, '12-6-2-13-7-7' and '12-6-2-16-7-2' can be related to '12-5-1-18-7-7' by crossover in the *D9S1677-D9S310* interval. There are at least 23 other chromosomes that would all predict crossovers in the *D9S1677-D9S310* interval but appear unrelated to the major three haplotypes except at *D9S1677*. Thus, between actual and historical crossovers, there are at minimum several crossover events that could move the distal border of the candidate region closer to *D9S1677* if additional markers are examined in the *D9S1677*-99E10 interval. We have searched this segment exhaustively for SSR polymorphisms (dinucleotide, trinucleotide and tetranucleotide) to no avail. However, we expect that our screening of candidate genes in this segment (see below) will reveal single base polymorphisms that can be useful in assessing this question.

On the centromeric side of *DYS*, there are two crossovers in the multicase families between *D9S748* and the genetic defect. In addition, there are five more recombination events between the next most centromeric marker *D9S172* and *DYS* that were not informative for *D9S748*. With additional markers, some of these five may prove to have occurred between *D9S748* and *DYS*. We expect that there will also be several historical recombination events in this interval because of the lack of strong disequilibrium between *DYS* and all proximal markers, including *D9S748*. Indeed, if the rate of decay in the conserved haplotype distal to *D9S1677* is mirrored on the proximal side, we would expect to detect historical crossovers within 100-200 kb centromeric to *D9S1677*.

Our current goal is to generate informative SSR polymorphisms in the vicinity of 13E8spc (Figure 1) at the centromeric side of the physical map (some 260 kb proximal to *D9S1677*) to determine whether these detect either the actual or historical crossovers. If a marker at this position remains identical in 98% of FD chromosomes, we will conclude that it is necessary to extend the contig to the centromeric side by additional YAC walking steps. On the other hand if the marker displays allelic differences on some founder FD chromosomes, it will be necessary to confirm that these changes reflect historical crossovers rather than mutation at the marker locus by isolating additional closely linked markers. Thus, if two such markers actually detect a historical crossover, it is to be expected that the two will form a consistent new haplotype on those chromosomes descendant from the recombinant individual. In practice, we also examine single base polymorphisms detected by SSCP or direct sequencing of PCR products to monitor the consistency of the haplotypes, as such sites are expected to have a lower mutation rate than SSR polymorphisms.

Of special note are the 3 haplotypes (representing 5 FD chromosomes) that appear at present to be unrelated to the ancestral founder haplotype even at *D9S1677*. These chromosomes could reflect independent mutations, or they could derive from the founder haplotype by crossovers proximal to *D9S1677*. Consequently, these chromosomes will also be examined for all new markers generated. If they become identical to the founder haplotype, increasing its frequency above 98% on FD chromosomes for several adjacent markers as we move proximally, they will infer new historical crossovers that exclude the region distal to *D9S1677*. If they remain unlike the founder haplotype even when a proximal limit is placed on the candidate region, this will provide strong evidence that they have independent mutations.

d.5 Exon amplification

Our first method for identifying genes in the candidate region is the technique of exon amplification developed and improved by Alan Buckler (27-30). This technique has the advantage that it is not dependent on abundance of an mRNA transcript as the genome is essentially the perfect "normalized" library. The sole disadvantage that we have found is that the technique cannot identify the small subset of genes with no internal exons, i.e. genes with no or only a single intron. We have considerable experience with exon amplification from positional cloning projects involving several disorders and from genome mapping projects on chromosomes 4, 9, 21 and 22. The method can be applied to individual cosmids or to pools of cosmids to reliably and rapidly isolate small single-copy coding sequence probes. Based on our work in the Huntington's disease (HD) region of chromosome 4, the false positive rate in identifying exons by this method is only ~10%. As this parameter is the most important determinant of how labor intensive subsequent analysis will be, it makes exon amplification a highly efficient method for finding genes, and essentially all multi-intron genes are identified. Less than 5% of the clones contain repeat sequences, and can be quickly discarded after sequencing and BLAST searching of GenBank. An additional 5% yield no signal on Northern blots and no cDNA clones with intensive library screening and are probably not authentic exons. The remaining 90% of exon clones yield cDNAs from library screening and most give positive results on Northern blots. As noted in the preliminary results, we have already isolated 75 exons (e.g., Figure 1) spread across the FD region that are used with other probes, STSs and SSRs to orient and overlap clones in the contigs. Typically, exon trapping of a single cosmid yields 2-10 exon products with a single cloning enzyme in the exon amplification step. As multiple cloning enzymes can be used to generate different exon-containing fragments, intensive application can yield even more exons. Indeed, in our search of the *PTPH1* region, we trapped 12 of the 24 internal exons in this gene.

Exon clones are typically 80-250 bp in size although larger clones can be generated if the cloning procedure in the exon amplification step is designed to capture multiple adjacent exons that are then correctly spliced together. The clones are rapidly sequenced (27-29) and all sequences are searched with BLASTX (for protein coding similarity in all frames) and BLASTN (for nucleotide sequence similarity) in GenBank. This step is designed to eliminate the few exons that contain repeat sequences and to identify those that are identical or closely related to known genes or motifs. The BLASTN program is also used to search the established databases (e.g. dbest) of expressed sequence tags for evidence of a known cDNA clone corresponding to the exon. Whether detected at this exon stage or at the cDNA stage (see next section), discovery of a related sequence in the public databases facilitates all further steps in the assembly and analysis of the gene sequence (e.g. our work on *PTPH1*).

The exon probes are then hybridized back to the cosmid/YAC contig to confirm that they derive from this segment and to establish their precise location. If there is any doubt about these hybridizations, the probe is also hybridized to Southern blots of DNA from the chromosome 9 containing somatic cell hybrids to confirm that the sequence derives from this autosome.

The exon probes are also hybridized to Northern blots containing RNA from a sampling of normal human tissues (Clontech). The results provide a clue as to the size of the corresponding mRNA and the most appropriate cDNA library for screening. They also indicate which neighboring exons probably represent different portions of an the same gene. Finally, the exon clones (and other probes such as end-clones, Alu-PCR products and cosmid fragments) are also hybridized to DNA blots of restriction digests comparing control individuals with FD individuals representing both major and minor haplotypes to search for evidence of a deletion or other rearrangement that might immediately pinpoint the site of the defect.

d.6 Isolation of candidate cDNAs

While isolation of exon clones identifies the presence of a gene in the region, we must isolate cDNA clones to define the full coding sequence and pattern of alternative splicing and expression. Fortunately, the exon clones are ideally suited to screening cDNA libraries by hybridization, either alone or in pools. We have found that almost all cDNAs, even those thought by Northern blot analysis to be limited to a peripheral tissue, can be isolated at some frequency from human fetal and adult brain cDNA libraries. Consequently, we standardly screen such libraries with new exons. Where the Northern blot analysis indicates different mRNA sizes or vastly different levels of expression in various tissues, we screen other cDNA libraries as needed to resolve the issues. For those exons that fail to yield cDNA clones immediately, we use a PCR assay to screen the fractions of a human brain cDNA library organized by on a pooling scheme that permits rapid identification of positive clones (67). Finally, we have begun to use high density filters from the IMAGE consortium (68) in an attempt to identify cDNAs in a single hybridization step without the need for purification by repeated rescreening.

The cDNA clones are compared by restriction digestion and selected clones are sequenced to assemble the entire sequence, discern any unspliced products and identify the open reading frame. Additional rounds of cDNA library screening or use of the RACE procedure may be needed at this stage to complete identification of the transcript particularly at the 5' end (63). Again, the sequences are subjected to BLAST analysis against GenBank, dbest, other EST databases and databases of sequence from lower organisms (especially *Saccharomyces cerevisiae* and *C. elegans*). Although identity with a sequence in the database can be detected at the exon level, detection of significant structural motifs or distantly related family members that might provide a clue to the gene's function are more likely to first occur at this stage. The cDNAs are mapped back to the cosmid/YAC contig to determine their location on the map and the extent of genomic coverage and Northern blots are performed to determine the pattern of expression. As with the exons, cDNAs are also used in Southern blot analysis to search for evidence of a genomic deletion in FD.

At this point, Northern blots with RNA from control and FD lymphoblasts are included to test for specific non-expression or altered mRNA size in FD. Although it could be argued that expression of *DYS* may be limited to a subset of neurons, cloning of genetic defects in other neurodegenerative disorders has usually identified a gene that is more widely expressed. Consequently, there is reason to believe that lymphoblasts will

express many of the genes detected. Unfortunately, there is too little FD autopsy tissue available to use Northern blot analysis of these as a primary screen. For those mRNAs not expressed in either control or FD lymphoblasts, RT-PCR screening of the FD autopsy tissues is possible.

When the cDNAs are placed on the physical map, we will note any evident gaps that might indicate the presence of genes that we have missed. First, we will target the specific exon trapping of the gap region with multiple restriction enzyme combinations to maximize our yield of potential exons. Failure to detect exons could reflect the presence of a gene with a single or no intron. If the gap persists, direct cDNA selection will be used to screen libraries (72). We have employed direct cDNA selection as a method for finding genes in cloned DNA and in particular have compared it with exon trapping in a 2 Mb cloned, sequenced region of human chromosome 4. Unlike exon trapping, direct selection can be biased for more abundantly expressed transcripts and it has a false positive rate (particularly with recovery of repetitive sequences) that makes it less favorable as a first line of attack. However, it represents an effective follow-up procedure, particularly as it can identify genes with one or no introns missed by exon trapping. As we will supply genomic clones to Dr. Eric Lander for sequencing in the MIT Genome Center, it may also be possible to target cDNA isolation in the gaps by sequence-based exon prediction. We have also compared exon trapping and exon prediction with the GRAIL2 program (73-75) on chromosome 4. We have found that the two techniques are complementary in that both identify many, but often different exons in a gene. In the *HD* gene, the combination of both methods found all but 12 of the 67 exons. The false positive rate in this region was about 50% for GRAIL2. Using the combination of techniques outlined above, we should be able to identify all of the genes in the FD candidate region.

d.6 Screening of candidate genes for mutation

The mutation in FD is expected to cause a loss of function of the *DYS* gene because of its recessive pattern of inheritance and the absence of any detectable abnormality in heterozygous FD carriers. If the disorder is caused by a subtle mutation, such as a single base change, then we will not detect a genomic deletion or altered expression of a gene in the FD candidate region. Such a mutation is most likely to affect the coding sequence of the mRNA, either directly by introducing an amino acid substitution, a stop codon or a frameshift, or indirectly by altering the splicing pattern. Consequently, RT-PCR offers a viable method for obtaining the most appropriate segment of the gene for mutation analysis. In our experience, most genes can be amplified by RT-PCR from lymphoblasts. For those that cannot, FD and control fibroblasts or autopsy tissues will be used. RT-PCR will be designed to amplify the coding sequence and proximal UTR sequences in large overlapping 2-3 kb segments. These can then be reamplified using internal primers to generate smaller PCR fragments for analysis. At this stage, evidence of any mutations that affect the splicing pattern will be detected as altered RT-PCR products and will lead us to sequence the intron regions adjacent to the corresponding exons.

If no genomic deletion or altered mRNA transcript is found, it will be necessary to examine the primary sequence of the RT-PCR products. There are many methods of scanning such segments for mutation, including SSCP, chemical cleavage, denaturing gradient gel electrophoresis, etc. We favor SSCP as a rapid scan that detects about 90% of alterations (56,76). These methods are fine for assembling a collection of mutations in a gene known to cause a genetic disorder but to be sure of identifying the crucial disease-producing mutation requires a more effective method. We have found that it is most efficient to simply use direct sequencing of the PCR products as part of the primary screen for mutations. Because we can use individuals homozygous for a particular haplotype, a change in the sequencing gel is readily detected, as evidenced in several polymorphisms that we have found.

Although the exact extent of the candidate region has yet to be defined, it will probably cover at most a few hundred kb (and perhaps far less) based on the number of historical and actual recombination events available. Thus, we expect that it will be necessary to examine 5-20 genes to identify the FD defect. The cDNAs that we find will be prioritized for analysis after considering a number of factors, including their location and extent of coverage of the candidate region, any function implied by the database searches that would make them better candidates, and expression pattern. At present, we have identified four potential genes centromeric to *D9S310* based on the clustered exons. two bear no relationship to known sequences in the public databases. One,

represented by seven trapped exons, is a previously unrecognized member of the protein 4.1 family of membrane-cytoskeleton associated proteins. The remaining gene shows weak similarity to a calcium channel.

When sequence differences are found in the candidate cDNAs, we will determine whether they are likely to represent polymorphisms or the pathogenic mutation. Clearly, any change that would result in truncation of the protein encoded at a locus (e.g., mutation to a stop codon or a frameshift causing premature termination) is likely to be an inactivating mutation. However, a missense substitution is harder to evaulate. We will assess the occurrence of any sequence change detcted on FD chromosome by genotyping Ashkenazi Jewish and non-Jewish control chromosomes, particularly those with surrounding marker alleles similar to the founder haplotype. If the change is not present on several hundred normal chromosomes but is on the founder FD haplotype, it will likely represent the genetic defect. In this case, we would expect that the rare FD haplotypes that are not ancestrally related to major founder haplotype would not show the change, but would exhibit another alteration in the same gene.

It is unlikely that after screening for genomic deletions, altered mRNA expression and altered coding sequence of all genes in the candidate region that we will not have identified the *DYS* gene. However, in this event the mutation is either in a gene that was missed or in a regulatory or other non-coding sequence. If at this point the FD region has been sequenced, it will be possible to design PCR strategies allowing comparison of the full genomic sequence between disease and normal chromosomes. An alternative strategy is to directly compare cloned normal and FD DNA from the candidate region by any of a number of techniques designed to find sequence differences (denaturing gradient gels, heteroduplex analysis, direct sequencing, etc.). The normal cosmid contig necessary for this approach will already be available from this work. However, we have also prepared a cosmid library from a patient homozygous for the founder FD haplotypes. Thus, if this unlikely worst-case scenario occurred, we could use the exon and cDNA probes from the candidate region to quickly assemble an equivalent cosmid contig representing the founder FD haplotype. Direct comparison of the control and FD genomic sequences would be expected to identify an average of one base pair difference/0.5-1 kb, each of which would then have to be examined to determine whether they represented a polymorphism. Any change present only on the founder haplotype and never seen on normal chromosomes would be a strong candidate for causing FD, but it would remain to design experiments aimed at identifying the functional consequence of such a change on adjacent expressed sequences.

d.7 Characterization of genetic defects in *DYS*

The FD defect could be found at any stage of our analysis, and is most likely to occur by identification of a genomic deletion, a splice alteration or an inactivating mutation occurring specifically on an FD chromosome and never on normal chromosomes. Part of the initial definition of *DYS* will be to examine the sequence of the culprit gene from all independent haplotypes associated with FD to define other mutations that implicate the same gene. However, we will also examine the gene in patients with other sensory neuropathies, particularly those with 'atypical' FD and those non-Jewish patients with related phenotypes. Identification of mutations on all the haplotypes of FD and perhaps in other neuropathies will permit genotype:phenotype correlations that might explain some of the wide variation in phenotypic expression and identify functionally important domains in the *DYS* protein.

From scanning of the public databases, particularly the EST databases and by low stringency screening of cDNA libraries, we will also determine whether the FD gene is unique, or a member of a multi-gene family. If related genes are found, these will also be mapped in the human genome and scanned for mutation in other neuropathies. Similarly, they may be useful reagents for comparison in the various biological and biochemical approaches necessary to define the normal function of *DYS*. The database scans will also reveal whether a homologue of *DYS* has been seen previously in any other mammalian species. On the basis of its map position in the human vis-à-vis flanking markers *ALDOB* and *HXB*, we expect that *DYS* will lie in a region of synteny conservation on mouse Chr 4 (77). We have mapped the mouse PTPH1 homologue to Chr 4 in support of this assumption. To date there are no mutations similar to FD reported to map to this region of the mouse genome.

However, the scanning of public databases for related sequences in lower organisms (e.g., yeast, *C. elegans*) will reveal whether there are homologues that map at or near sites of known mutations in these genetic models. These would also facilitate delineation of the normal physiological function of *DYS* and present an alternative more tractable system for studying functional and genetic interactions.

d.8 Characterization of *DYS*-Structure and expression

When the FD defect has been found, it will target the function of a particular gene as being crucial for development and maintenance of sensory and autonomic neurons. However, the experience in other neurologic disorders has been that defects that cause specific neuronal deficits usually occur in genes whose expression is not limited to the target cells but is more widespread. It is possible, but unlikely, that the *DYS* gene will be one that has already been reported and for which basic information concerning structure and expression pattern are known. More likely, the *DYS* gene will prove to encode either a novel protein or a novel member of a known family of proteins. In either case, we will determine its genomic structure and pattern of expression. We expect, as was the case with *PTPH1*, that much of the genomic structure will fall out of the exon trapping analysis and mutation scanning. The remainder will be elucidated by direct sequencing of cosmids using primers based on the cDNA.

Similarly, we expect that the RT-PCR procedure will have defined any common splice variants expressed in lymphoblasts or in the target tissue used for mutation screening. However, we will assess the expression of the gene in other human fetal and adult tissues by Northern blot analysis and, where meaningful, by RT-PCR to detect novel splice variants. As the *DYS* gene is likely to be expressed throughout development of the nervous system, we will assess its pattern of expression in mouse development and, if possible in human development, by *in situ* hybridization to mRNA (78,79).

d.9 Characterization of pathogenesis due to mutation in *DYS*

Once a basic understanding of the sequence and expression pattern of the *DYS* gene and of the nature of the FD mutation has been achieved, the crucial questions to be resolved will surround its mode of action in causing sensory and autonomic neuropathy. The directions that can be taken are many and varied, but will depend on the nature of *DYS*. If it is a known gene, or a member of a known family of genes, then many of the experimental systems necessary to study the effects of the mutation will already have been established by others and we can collaborate with investigators expert in those systems. For example, if the *DYS* gene proves to be a previously unsuspected member of the neurotrophin family of genes then consideration of its expression in the experimental paradigms used to differentiate the roles of NGF, BNDF, NT-3, NT-4, GDNF, etc., would be called for (37). However, in the event that *DYS* encodes a novel gene, we will have to develop the reagents necessary for exploring its biochemical activity and physiological role.

The directions that can be taken to investigate the normal physiological role of the *DYS* gene product and the pathogenic consequences of its failure in FD can be broadly described in three complementary (and overlapping) categories: biochemical, cell biological and genetic.

1. Biochemical approaches

An obvious and important step that will be required if *DYS* encodes a novel gene product will be the preparation of specific antisera to detect the protein. These will be raised in rabbits (East Acres Biologicals) against fusion proteins expressed in *E. coli* (we typically use GST fusions) and against synthetic peptides if it is necessary to choose a small unique segment that distinguishes the protein from a related family member. The antisera can then be used in Western blot analyses to examine the size of the protein, its pattern of expression in tissues, potential processing or modifications, regulation and its localization in cell fractionation studies. Similarly immunoprecipitation experiments can be used to monitor the protein's stability, processing and/or breakdown and binding with other proteins. The antisera can also be used for immunohistochemical studies to examine the cellular and subcellular distribution of the protein in human (and rodent) tissues. Another step in characterization might be to purify the protein, either from tissues or from baculovirus expression from tagged constructs, and to determine whether it is present in cells as a monomer or complexed with other proteins. A

primary strategy that has recently emerged to study novel proteins is to define their pattern of interaction with other proteins in the cell using so-called "two-hybrid" systems. We have used primarily the protein interaction trap system developed by the Roger Brent lab at MGH (80,81), the 'Chevray' system (82) and a modified system developed by Mark Vidal in Ed Harlow's lab at MGH (M. Vidal, pers. commun.). The latter has numerous advantages in reducing the false positive rate and allowing the strength of interaction to be titrated. Depending on the nature of the *DYS* protein, the two-hybrid system may be a useful way of obtaining a clue to its normal function.

b. Cell biological approaches

Cell biological approaches to a novel protein can provide clues to its subcellular localization, interactions and regulation. Development of epitope-tag expression constructs can permit the overexpression of the protein in various cell types to monitor for biological effects and for cellular localization. The tagged proteins can also be used to test for *in vivo* interaction with candidate interacting proteins identified by two-hybrid analysis or can aid in direct purification of the tagged product and associated proteins. Depending on the expression of the *DYS* protein it might be tested in experimental paradigms designed to monitor survival or death of primary sensory or autonomic neurons, or their response to various manipulations. Antisense experiments to reduce expression of the protein can be similarly be used to measure its biological effects in these systems. Ultimately, one would like to develop a reliable tissue culture system in which the clear effects of deficiency of the *DYS* protein parallel some aspect of the pathology, and are rescuable by reintroduction of the protein.

c. Genetic approaches

Perhaps the most powerful approach in the long-term is relating the function of the *DYS* gene in man to homologues in lower organisms where more powerful genetic analysis can be performed. By simply comparing the sequence of the *DYS* gene and its homologues in mouse and other organisms, it may be possible to target functional domains in the protein product that have been under strict evolutionary constraints.

We will also sequence and map the mouse cDNA to confirm that the *DYS* homologue maps on Chr 4. The information on the mouse sequence and studies of expression of the *DYS* homologue in mouse develop will act as the foundation for the creation of a genetic model in the mouse. The goal would be to create a mouse "knock-out" similar to that which we generated for the mouse Huntington's disease homologue, *Hdh* (83). The hope is that a *DYS* knock-out would develop a sensory and autonomic neuropathy comparable to FD providing a manipulable experimental system in which to carry out further biochemical and cell biological studies. This would also permit the an assessment of whether the significant variation in manifestation of the disorder in FD patients with the classic founder haplotype reflects the operation of modifier genes that can be mapped using mouse crosses. Identification of such modifying loci, particularly those that seem to allow milder expression of the disorder, might be one of the best routes to understanding the pathogenesis of the disease and to providing a treatment. Generation of a mouse *DYS* knock-out would also permit this model to be compared with the several mouse lines with inactivated neurotrophins or neurotrophin receptors (36,37)]. Similarly, it would be interesting to explore expression of the *DYS* gene in these other knock-out animals .

If there are recognizable homologues of *DYS* in *Drosophila, C. elegans* or yeast, these too offer the potential to develop a genetic model system in which inactivation of the gene function can be related directly to a phenotype. While that phenotype is less likely to show a parallel with that of the human disorder, the fundamental biochemical processes underlying it are probably closely related. Consequently, identifying suppressors and enhancers of the phenotype would provide genetic interactors that might identify the biochemical or functional pathway in which *DYS*'s normal physiological role resides.

d.10 Gender and Minority Inclusion

The ethnic bias of the FD defect dictates that all of the human subjects studied in the initial stage of this research plan will be of Ashkenazi Jewish descent. Once the gene has been found, the research will extend to other patients with hereditary sensory neuropathies who may be of any race or ethnic background. As there is no gender bias in FD, the study group will consist equally of males and females.

d.11 Conclusion

The preliminary work that we have done places us in a strong position to identify the FD gene within the term of this grant. Doing so will identify a protein whose activity is crucial to the development and survival of sensory and autonomic neurons. While this protein could be a novel growth factor or receptor, the results of other positional cloning projects have shown that it cannot be predicted in advance what the nature of the *DYS* protein will be. If the gene defect that is identified implicates an obvious protein function, then it will be straightforward to turn to those investigators and experimental systems that can help explore that function and relate it to FD pathogenesis. If the protein is novel, a more complex approach will be required as more basic information and reagents will have to be developed to form educated hypotheses concerning the protein's potential function. Which experimental strategies are best will depend on whether the *DYS* protein is likely to reside in the membrane, cytoplasm, nucleus or in an organelle, whether it has functional motifs, whether it shows restricted or widespread expression and many other factors. We have outlined some of the potential directions that could be pursued, but clearly each of these constitutes a significant effort and it is not possible to commit to all of them in advance. Consequently, we will direct our focus to isolating and characterizing the FD defect, with a knowledge that once the gene is found, we can be guided in investigating the mode of FD pathogenesis by the nature of the *DYS* gene product and by our experience in pursuing similar gene characterization projects in Huntington's disease, neurofibromatosis 2, and Batten's disease. The product of our work will be a knowledge of the cause of FD and its relationship to other sensory neuropathies, the ability to screen for carriers of the disorder, and a beginning to the exploration of the normal and abnormal function of the DYS gene in human and animal model systems. In the long-term we will contribute both to the fundamental understanding of development and maintenance of the sensory and autonomic nervous systems, and to the hope of FD patients for an effective treatment preventing progression of this devastating disorder.

e. Human Subjects

Human subjects for this study will consist of FD affected children and adults, their parents and unaffected siblings. In addition, individuals affected with other sensory neuropathies, their parents and siblings will be collected where possible. Material collected will consist of blood samples and autopsy tissues. Most of the samples will be provided to us by clinical collaborators without direct contact between the MGH personnel and the human subjects. These collaborators will explain the nature of the research project and seek appropriate documented consent through an IRB-approved consent form. We anticipate collecting ~200 blood samples from this group to augment the existing ~500 cell lines and DNA samples obtained previously from our clinical collaborators in NY and Israel who are in contact with the majority of the FD families. There are only minimal risks associated with blood drawing. All human tissue samples collected are specifically designated for research. Blood samples are received in the laboratory and coded for confidentiality.

f. Vertebrate Animals

We anticipate no use of vertebrate animals in this project until the FD gene has been identified. At that point, specific research strategies that employ vertebrate animals may be advantageous, and we will seek institutional approval of specific protocols and clear these with the NIH before embarking on such studies.

g. Literature Cited

1. Axelrod, F. B., Nachtigal, R., and Dancis, J., 1974, Familial dysautonomia: diagnosis, pathogenesis and management, *Adv Pediatr* **21:**75-96.
2. Axelrod, F. B., 1984, Familial dysautonomia and other congenital and sensory autonomic neuropathies, in: *Cell and Molecular Biology of Neuronal Development* (I. B. Blake, ed.), Plenum Press, New York, pp. 331-340.
3. Riley, C. M., Day, R. L., Greely, D., and Langford, W. S., 1949, Central autonomic dysfunction with defective lacrimation, *Pediatrics* **3:**468-477.
4. Axelrod, F. B., and Pearson, J., 1984, Congenital sensory neuropathies. Diagnostic distinction from familial dysautonomia, *Am J Dis Child* **138:**947-54.
5. Brunt, P. W., and McKusick, V. A., 1970, Familial dysautonomia. A report of genetic and clinical studies, with a review of the literature, *Medicine (Baltimore)* **49:**343-74.
6. Axelrod, F. B., Porges, R. F., and Sein, M. E., 1987, Neonatal recognition of familial dysautonomia, *J Pediatr* **110:**946-8.
7. Mahloudji, M., Brunt, P. W., and McKusick, V. A., 1970, Clinical neurological aspects of familial dysautonomia, *J Neurol Sci* **11:**383-95.
8. Pearson, J., 1979, Familial dysautonomia (a brief review), *J Auton Nerv Syst* **1:**119-26.
9. Pearson, J., and Pytel, B., 1978, Quantitative studies of ciliary and sphenopalatine ganglia in familial dysautonomia, *J Neurol Sci* **39:**123-30.
10. Pearson, J., and Pytel, B. A., 1978, Quantitative studies of sympathetic ganglia and spinal cord intermediolateral gray columns in familial dysautonomia, *J Neurol Sci* **39:**47-59.
11. Pearson, J., Pytel, B. A., Grover-Johnson, N., Axelrod, F., and Dancis, J., 1978, Quantitative studies of dorsal root ganglia and neuropathologic observations on spinal cords in familial dysautonomia, *J Neurol Sci* **35:**77-92.
12. Axelrod, F. B., 1995, Familial dysautonomia, in: *Disorders of the Autonomic Nervous System* (D. Robertson, and I. Biaggioni, eds.), Harwood Academic Publishers, Luxembourg, pp. 217-231.
13. Breakefield, X. O., Orloff, G., Castiglione, C., Coussens, L., Axelrod, F. B., and Ullrich, A., 1984, Structural gene for beta-nerve growth factor not defective in familial dysautonomia, *Proc Natl Acad Sci U S A* **81:**4213-6.
14. Schwartz, J. P., and Breakefield, X. O., 1980, Altered nerve growth factor in fibroblasts from patients with familial dysautonomia, *Proc Natl Acad Sci U S A* **77:**1154-8.
15. Wrathall, J. R., 1986, Reduced neuronotrophic activity of fibroblasts from individuals with dysautonomia in cultures of newborn mouse sensory ganglion cells, *Brain Res* **364:**23-9.
16. Breakefield, X. O., Ozelius, L., Bothwell, M. A., Chao, M. V., Axelrod, F., Kramer, P. L., Kidd, K. K., Lanahan, A. A., Johnson, D. E., Ross, A. H., and et al., 1986, DNA polymorphisms for the nerve growth factor receptor gene exclude its role in familial dysautonomia, *Mol Biol Med* **3:**483-94.
17. Axelrod, F. B., Golstein, D. S., Holmes, C., Berlin, D., and Kopin, I. J., 1996, Pattern of plasma levels of catechols in familial dysautonomia, *Clinical Autonomic Research* :in press.
18. Pereira, J., Boyd, B., Newbigging, J., Lingwood, C., and Strasberg, P. M., 1994, Increased globotriaosylceramide on plasma membranes of synchronized familial dysautonomia cells. Verotoxin binding studies, *J Mol Neurosci* **5:**121-32.
19. Strasberg, P., Yeger, H., and Warren, I., 1992, Increased globotriaosylceramide in familial dysautonomia, *Lipids* **27:**978-83.
20. McKusick, V. A., 1966, Clinical genetics at a populational level. The ethnicity of disease in the United States, *Ala J Med Sci* **3:**408-24.
21. Moses, S. W., Rotem, Y., Jagoda, N., Talmor, N., Eichhorn, F., and Levin, S., 1967, A clinical, genetic and biochemical study of familial dysautonomia in Israel, *Isr J Med Sci* **3:**358-71.
22. Maayan, C., Kaplan, E., Shachar, S., Peleg, O., and Godfrey, S., 1987, Incidence of familial dysautonomia in Israel 1977-1981, *Clin Genet* **32:**106-8.

23. Blumenfeld, A., Axelrod, F. B., Trofatter, J. A., Maayan, C., Lucente, D. E., Slaugenhaupt, S. A., Liebert, C. B., Ozelius, L. J., Haines, J. L., Breakefield, X. O., and et al., 1993, Exclusion of familial dysautonomia from more than 60% of the genome, *J Med Genet* **30:**47-52.
24. Blumenfeld, A., Slaugenhaupt, S. A., Axelrod, F. B., Lucente, D. E., Maayan, C., Liebert, C. B., Ozelius, L. J., Trofatter, J. A., Haines, J. L., Breakefield, X. O., and et al., 1993, Localization of the gene for familial dysautonomia on chromosome 9 and definition of DNA markers for genetic diagnosis, *Nat Genet* **4:**160-4.
25. Eng, C. M., Slaugenhaupt, S. A., Blumenfeld, A., Axelrod, F. B., Gusella, J. F., and Desnick, R. J., 1995, Prenatal diagnosis of familial dysautonomia by analysis of linked CA- repeat polymorphisms on chromosome 9q31-q33, *Am J Med Genet* **59:**349-55.
26. Oddoux, C., Reich, E., Axelrod, F., Blumenfeld, A., Maayan, C., Slaugenhaupt, S., Gusella, J., and Ostrer, H., 1995, Prenatal diagnostic testing for familial dysautonomia using linked genetic markers, *Prenat Diagn* **15:**817-26.
27. Trofatter, J. A., Long, K. R., Murrell, J. R., Stotler, C. J., Gusella, J. F., and Buckler, A. J., 1995, An expression-independent catalog of genes from human chromosome 22, *Genome Res* **5:**214-24.
28. Church, D. M., Stotler, C. J., Rutter, J. L., Murrell, J. R., Trofatter, J. A., and Buckler, A. J., 1994, Isolation of genes from complex sources of mammalian genomic DNA using exon amplification, *Nat Genet* **6:**98-105.
29. Church, D. M., Banks, L. T., Rogers, A. C., Graw, S. L., Housman, D. E., Gusella, J. F., and Buckler, A. J., 1993, Identification of human chromosome 9 specific genes using exon amplification, *Hum Mol Genet* **2:**1915-20.
30. Buckler, A. J., Chang, D. D., Graw, S. L., Brook, J. D., Haber, D. A., Sharp, P. A., and Housman, D. E., 1991, Exon amplification: a strategy to isolate mammalian genes based on RNA splicing, *Proc Natl Acad Sci U S A* **88:**4005-9.
31. Collins, F. S., 1995, Positional cloning moves from perditional to traditional. *Nat Genet* **9:**347-50.
32. Huntington's Disease Collaborative Research Group, 1993, A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes., *Cell* **72:**971-83.
33. Trofatter, J. A., MacCollin, M. M., Rutter, J. L., Murrell, J. R., Duyao, M. P., Parry, D. M., Eldridge, R., Kley, N., Menon, A. G., Pulaski, K., and et al., 1993, A novel moesin-, ezrin-, radixin-like gene is a candidate for the neurofibromatosis 2 tumor suppressor [published erratum appears in Cell 1993 Nov 19;75(4):826], *Cell* **72:**791-800.
34. International Batten Disease Consortium, 1995, Isolation of a novel gene underlying Batten disease, CLN3. The International Batten Disease Consortium, *Cell* **82:**949-57.
35. Barde, Y. A., 1994, Neurotrophins: a family of proteins supporting the survival of neurons, *Prog Clin Biol Res* **390:**45-56.
36. Johnson, J., and Oppenheim, R., 1994, Neurotrophins. Keeping track of changing neurotrophic theory, *Curr Biol* **4:**662-5.
37. Snider, W. D., 1994, Functions of the neurotrophins during nervous system development: what the knockouts are teaching us, *Cell* **77:**627-38.
38. Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., Watkins, P. C., Ottina, K., Wallace, M. R., Sakaguchi, A. Y., and et al., 1983, A polymorphic DNA marker genetically linked to Huntington's disease, *Nature* **306:**234-8.
39. Seizinger, B. R., Rouleau, G. A., Ozelius, L. J., Lane, A. H., Faryniarz, A. G., Chao, M. V., Huson, S., Korf, B. R., Parry, D. M., Pericak-Vance, M. A., and et al., 1987, Genetic linkage of von Recklinghausen neurofibromatosis to the nerve growth factor receptor gene, *Cell* **49:**589-94.
40. Rouleau, G. A., Wertelecki, W., Haines, J. L., Hobbs, W. J., Trofatter, J. A., Seizinger, B. R., Martuza, R. L., Superneau, D. W., Conneally, P. M., and Gusella, J. F., 1987, Genetic linkage of bilateral acoustic neurofibromatosis to a DNA marker on chromosome 22, *Nature* **329:**246-8.

41. Seizinger, B. R., Rouleau, G. A., Ozelius, L. J., Lane, A. H., Farmer, G. E., Lamiell, J. M., Haines, J., Yuen, J. W., Collins, D., Majoor-Krakauer, D., and et al., 1988, Von Hippel-Lindau disease maps to the region of chromosome 3 associated with renal cell carcinoma, *Nature* **332:**268-9.
42. St George-Hyslop, P., Haines, J., Rogaev, E., Mortilla, M., Vaula, G., Pericak-Vance, M., Foncin, J. F., Montesi, M., Bruni, A., Sorbi, S., and et al., 1992, Genetic evidence for a novel familial Alzheimer's disease locus on chromosome 14, *Nat Genet* **2:**330-4.
43. Ozelius, L., Kramer, P. L., Moskowitz, C. B., Kwiatkowski, D. J., Brin, M. F., Bressman, S. B., Schuback, D. E., Falk, C. T., Risch, N., de Leon, D., and et al., 1989, Human gene for torsion dystonia located on chromosome 9q32-q34, *Neuron* **2:**1427-34.
44. Fontaine, B., Khurana, T. S., Hoffman, E. P., Bruns, G. A., Haines, J. L., Trofatter, J. A., Hanson, M. P., Rich, J., McFarlane, H., Yasek, D. M., and et al., 1990, Hyperkalemic periodic paralysis and the adult muscle sodium channel alpha-subunit gene, *Science* **250:**1000-2.
45. Velinov, M., Slaugenhaupt, S. A., Stoilov, I., Scott, C. I., Jr., Gusella, J. F., and Tsipouras, P., 1994, The gene for achondroplasia maps to the telomeric region of chromosome 4p [see comments], *Nat Genet* **6:**314-7.
46. Tanzi, R. E., Petrukhin, K., Chernov, I., Pellequer, J. L., Wasco, W., Ross, B., Romano, D. M., Parano, E., Pavone, L., Brzustowicz, L. M., and et al., 1993, The Wilson disease gene is a copper transporting ATPase with homology to the Menkes disease gene, *Nat Genet* **5:**344-50.
47. McClatchey, A. I., Van den Bergh, P., Pericak-Vance, M. A., Raskind, W., Verellen, C., McKenna-Yasek, D., Rao, K., Haines, J. L., Bird, T., Brown, R. H., Jr., and et al., 1992, Temperature-sensitive mutations in the III-IV cytoplasmic loop region of the skeletal muscle sodium channel gene in paramyotonia congenita, *Cell* **68:**769-74.
48. Jones, C., Kao, F. T., and Taylor, R. T., 1980, Chromosomal assignment of the gene for folylpolyglutamate synthetase to human chromosome 9, *Cytogenet Cell Genet* **28:**181-94.
49. Warburton, D., Gersen, S., Yu, M. T., Jackson, C., Handelin, B., and Housman, D., 1990, Monochromosomal rodent-human hybrids from microcell fusion of human lymphoblastoid cells containing an inserted dominant selectable marker, *Genomics* **6:**358-66.
50. McCormick, M. K., Buckler, A., Bruno, W., Campbell, E., Shera, K., Torney, D., Deaven, L., and Moyzis, R., 1993, Construction and characterization of a YAC library with a low frequency of chimeric clones from flow-sorted human chromosome 9, *Genomics* **18:**553-8.
51. Haluska, F. G., Fountain, J. W., Gitai, Z., Rivera, M., Stanton, V. P., and Housman, D. E., 1994, Isolation of YAC contigs spanning overlapping chromosome 9p21 deletions in malignant melanoma (Meeting abstract), *Proc Annu Meet Am Assoc Cancer Res* **35:**A1125.
52. Hunter, K. W., Ontiveros, S. D., Watson, M. L., Stanton, V. P., Jr., Gutierrez, P., Bhat, D., Rochelle, J., Graw, S., Ton, C., Schalling, M., and et al., 1994, Rapid and efficient construction of yeast artificial chromosome contigs in the mouse genome with interspersed repetitive sequence PCR (IRS-PCR): generation of a 5-cM, > 5 megabase contig on mouse chromosome 1, *Mamm Genome* **5:**597-607.
53. Liu, J., Stanton, V. P., Jr., Fujiwara, T. M., Wang, J. X., Rezonzew, R., Crumley, M. J., Morgan, K., Gros, P., Housman, D., and Schurr, E., 1995, Large-scale cloning of human chromosome 2-specific yeast artificial chromosomes (YACs) using an interspersed repetitive sequences (IRS)- PCR approach, *Genomics* **26:**178-91.
54. Moir, D. T., Dorman, T. E., Xue, F., Ma, N. S., Stanton, V. P., Jr., Housman, D., Bowden, D. W., Noll, W. W., and Mao, J., 1993, Rapid identification of overlapping YACs in the MEN2 region of human chromosome 10 by hybridization with Alu element-mediated PCR products, *Gene* **136:**177-83.
55. Albertsen, H. M., Abderrahim, H., Cann, H. M., Dausset, J., Le Paslier, D., and Cohen, D., 1990, Construction and characterization of a yeast artificial chromosome library containing seven haploid human genome equivalents, *Proc Natl Acad Sci U S A* **87:**4256-60.

56. Povey, S., Armour, J., Farndon, P., Haines, J. L., Knowles, M., Olopade, F., Pilz, A., White, J. A., and Kwiatkowski, D. J., 1994, Report and abstracts of the Third International Workshop on Chromosome 9. Cambridge, United Kingdom, 9-11 April, 1994, *Ann Hum Genet* **58:**177-250.
57. Deaven, L. L., Van Dilla, M. A., Bartholdi, M. F., Carrano, A. V., Cram, L. S., Fuscoe, J. C., Gray, J. W., Hildebrand, C. E., Moyzis, R. K., and Perlman, J., 1986, Construction of human chromosome-specific DNA libraries from flow- sorted chromosomes, *Cold Spring Harb Symp Quant Biol* **51 Pt 1:**159-67.
58. Van Dilla, M. A., and Deaven, L. L., 1990, Construction of gene libraries for each human chromosome, *Cytometry* **11:**208-18.
59. Shero, J. H., McCormick, M. K., Antonarakis, S. E., and Hieter, P., 1991, Yeast artificial chromosome vectors for efficient clone manipulation and mapping, *Genomics* **10:**505-8.
60. Lucente, D., Chen, H. M., Shea, D., Samec, S. N., Rutter, M., Chrast, R., Rossier, C., Buckler, A., Antonarakis, S. E., and McCormick, M. K., 1995, Localization of 102 exons to a 2.5 Mb region involved in Down syndrome, *Hum Mol Genet* **4:**1305-11.
61. Murrell, J., Trofatter, J., Rutter, M., Cutone, S., Stotler, C., Rutter, J., Long, K., Turner, A., Deaven, L., Buckler, A., and et al., 1995, A 500-kilobase region containing the tuberous sclerosis locus (TSC1) in a 1.7-megabase YAC and cosmid contig, *Genomics* **25:**59-65.
62. Yang, Q., and Tonks, N. K., 1991, Isolation of a cDNA clone encoding a human protein-tyrosine phosphatase with homology to the cytoskeletal-associated proteins band 4.1, ezrin, and talin, *Proc Natl Acad Sci U S A* **88:**5949-53.
63. Frohman, M. A., Dush, M. K., and Martin, G. R., 1988, Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer, *Proc Natl Acad Sci U S A* **85:**8998-9002.
64. Zahn, L. M., and Kwiatkowski, D. J., 1995, A 37-marker PCR-based genetic linkage map of human chromosome 9: observations on mutations and positive interference, *Genomics* **28:**140-6.
65. Weissenbach, J., Gyapay, G., Dib, C., Vignal, A., Morissette, J., Millasseau, P., Vaysseix, G., and Lathrop, M., 1992, A second-generation linkage map of the human genome [see comments], *Nature* **359:**794-801.
66. Dib, C., Faure, S., Fizames, C., Samson, D., Drouot, N., Vignal, A., Millasseau, P., Marc, S., Hazan, J., Seboun, E., Lathrop, M., Gyapay, G., Morissette, J., and Weissenbach, J., 1996, A comprehensive genetic map of the human genome based on 5,264 microsatellites [see comments], *Nature* **380:**152-4.
67. Munroe, D. J., Loebbert, R., Bric, E., Whitton, T., Prawitt, D., Vu, D., Buckler, A., Winterpacht, A., Zabel, B., and Housman, D. E., 1995, Systematic screening of an arrayed cDNA library by PCR, *Proc Natl Acad Sci U S A* **92:**2209-13.
68. Lennon, G., Auffray, C., Polymeropolous, M., and Soares, M. B., 1996, The I.M.A.G.E. Consortium: An integrate molecular analysis of genomes and their expression, *Genomics* **33:**151-152.
69. McCabe, E. R., 1991, Utility of PCR for DNA analysis from dried blood spots on filter paper blotters, *PCR Methods Appl* **1:**99-106.
70. Anderson, M. A., and Gusella, J. F., 1984, Use of cyclosporin A in establishing Epstein-Barr virus-transformed human lymphoblastoid cell lines, *In Vitro* **20:**856-8.
71. Baxendale, S., Bates, G. P., MacDonald, M. E., Gusella, J. F., and Lehrach, H., 1991, The direct screening of cosmid libraries with YAC clones, *Nucleic Acids Res* **19:**6651.
72. Lovett, M., 1994, Fishing for complements: finding genes by direct selection, *Trends Genet* **10:**352-7.
73. Xu, Y., Mural, R. J., and Uberbacher, E. C., 1994, Constructing gene models from accurately predicted exons: an application of dynamic programming, *Comput Appl Biosci* **10:**613-23.
74. Xu, Y., Mural, R., Shah, M., and Uberbacher, E., 1994, Recognizing exons in genomic sequence using GRAIL II, *Genet Eng (N Y)* **16:**241-53.
75. Xu, Y., Einstein, J. R., Mural, R. J., Shah, M., and Uberbacher, E. C., 1994, An improved system for exon recognition and gene modeling in human DNA sequences, *Ismb* **2:**376-84.

76. MacCollin, M., Ramesh, V., Jacoby, L. B., Louis, D. N., Rubio, M. P., Pulaski, K., Trofatter, J. A., Short, M. P., Bove, C., Eldridge, R., and et al., 1994, Mutational analysis of patients with neurofibromatosis 2, *Am J Hum Genet* **55:**314-20.
77. Pericak-Vance, M. A., Bale, A. E., Haines, J. L., Kwiatkowski, D. J., Pilz, A., Slaugenhaupt, S., White, J. A., Edwards, J., Marchuk, D., and Povey, S., 1995, Report on the Fourth International Workshop on Chromosome 9, *Annal Hum Genet* **59**.
78. Landwehrmeyer, G. B., McNeil, S. M., Dure, L. S. t., Ge, P., Aizawa, H., Huang, Q., Ambrose, C. M., Duyao, M. P., Bird, E. D., Bonilla, E., and et al., 1995, Huntington's disease gene: regional and cellular expression in brain of normal and affected individuals, *Ann Neurol* **37:**218-30.
79. Dure, L. S. t., Landwehrmeyer, G. B., Golden, J., McNeil, S. M., Ge, P., Aizawa, H., Huang, Q., Ambrose, C. M., Duyao, M. P., Bird, E. D., and et al., 1994, IT15 gene expression in fetal human brain, *Brain Res* **659:**33-41.
80. Gyuris, J., Golemis, E., Chertkov, H., and Brent, R., 1993, Cdi1, a Human G1 and S phase protein phosphatase that associates with Cdk2., *Cell* **75:**791-803.
81. Zervos, A. S., Gyuris, J., and Brent, R., 1993, Mxi1, a protein that specifically interacts with Max to bind Myc-Max recognition sites., *Cell* **72:**223-232.
82. Chevray, P. M., and Nathans, D., 1992, Protein interaction cloning in yeast: identification of mammalian proteins that react with the leucine zipper of Jun, *Proc Natl Acad Sci U S A* **89:**5789-93.
83. Duyao, M. P., Auerbach, A. B., Ryan, A., Persichetti, F., Barnes, G. T., McNeil, S. M., Ge, P., Vonsattel, J. P., Gusella, J. F., Joyner, A. L., and et al., 1995, Inactivation of the mouse Huntington's disease gene homolog Hdh, *Science* **269:**407-10.

h. Consortium/Contractual Arrangements

NONE

i. Consultants

Felicia Axelrod. M.D. (NYU) will serve as our primary clinical collaborator (see attached letter and biosketch) along with Drs. Channa Mayaan and Anat Blumenfeld in Israel (see attached letters).

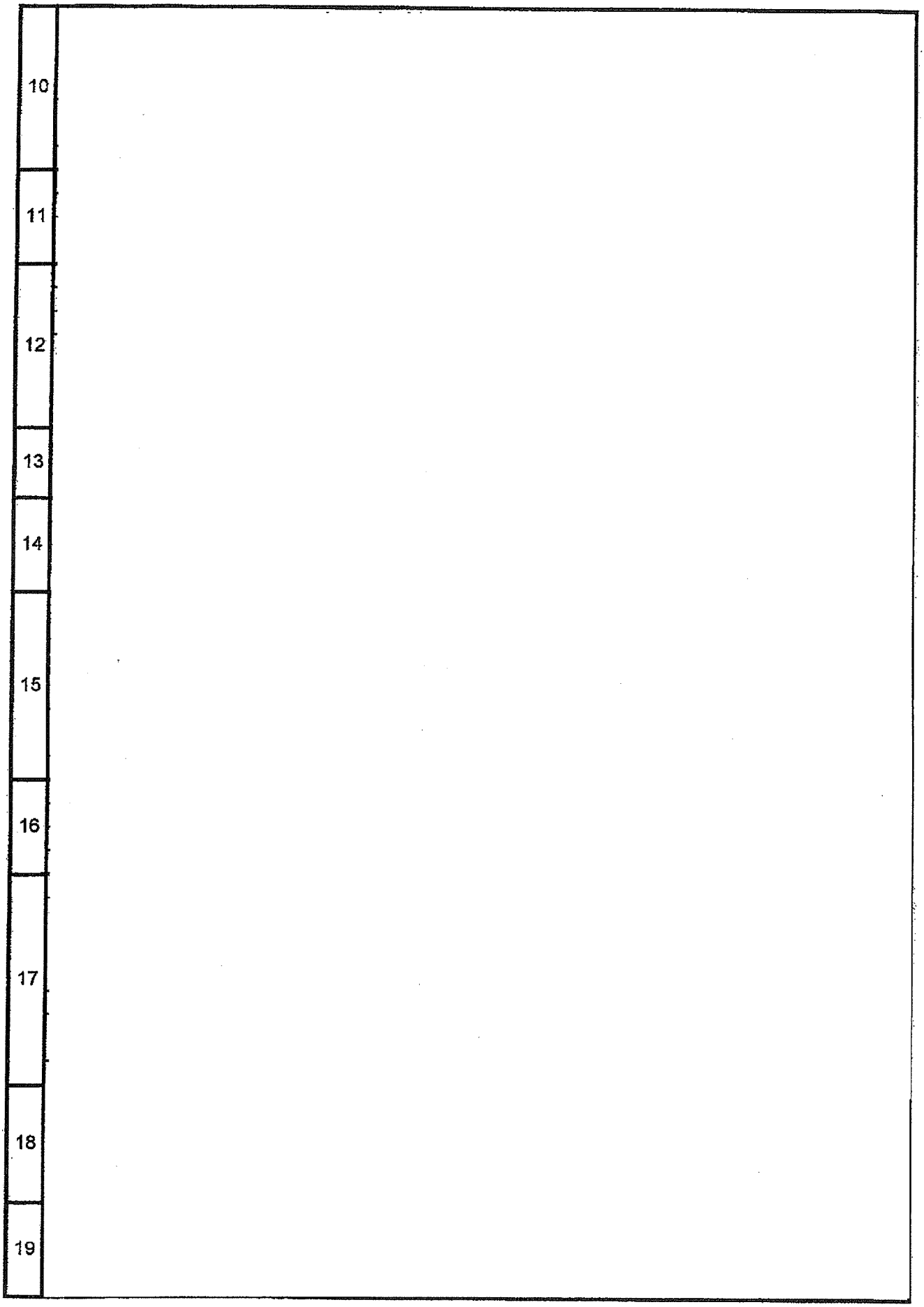
EXHIBIT 2

pcrCOSMID - DYS129

	PCR Product	Results
1		
2		
3		
4		
5		
6		
7		
8		
9		

REDACTED

pcrCOSMID - DYS129



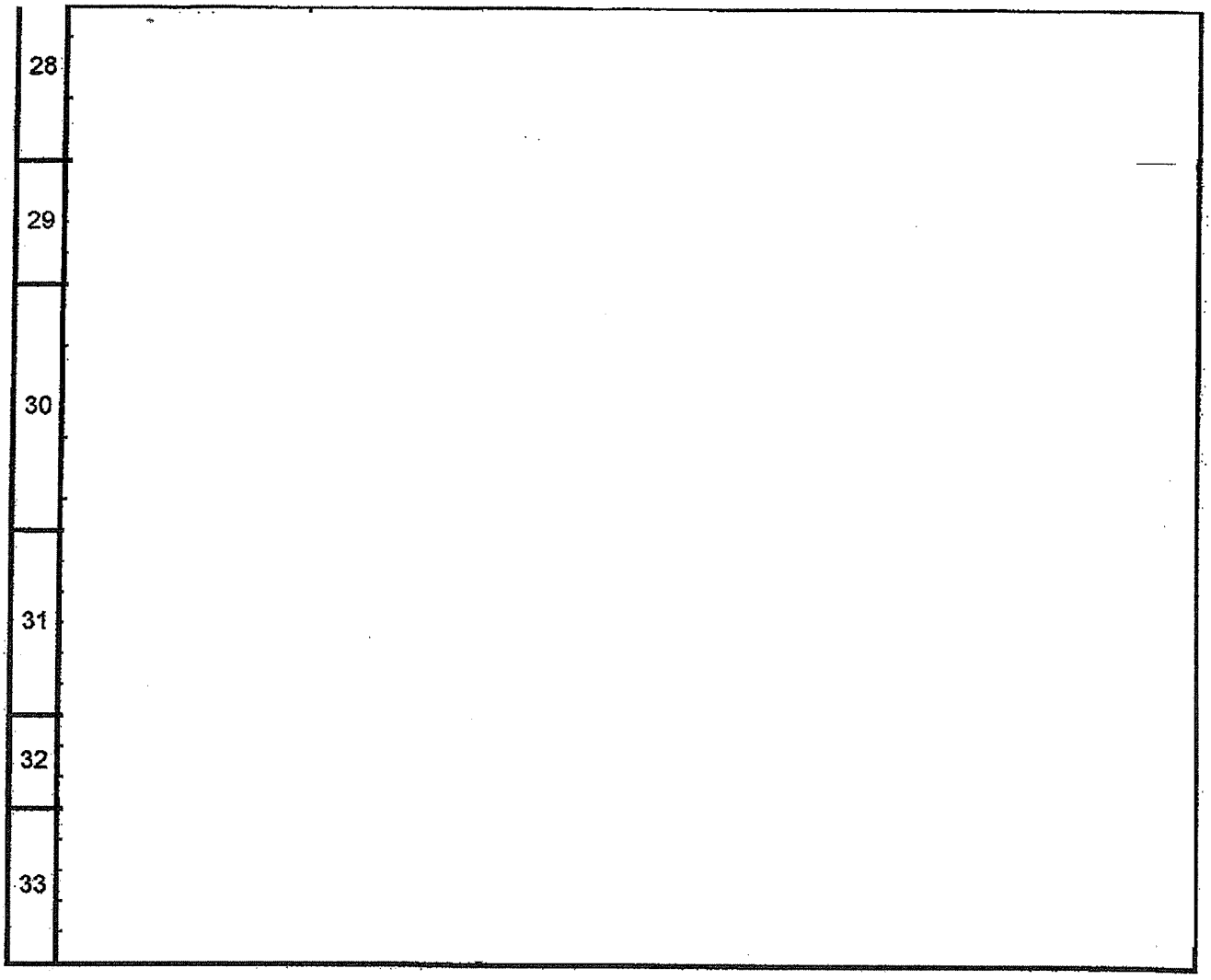
REDACTED

pcrCOSMID - DYS129

20		
21		
22		
23		
24		
25	2003	NOT done - poly - bp 76200, stretch of T's but DYS129 has one extra, genpoly 129. The primer worked a little better in control cosmid but not much - re-do. Gap between 2003 and 2004.
	2004	done - doesn't get past a stretch of T's though..
	2005	done - didn't work at all in control cosmid and after 80 bp's it hits the stretch of T's and dies. This stretch of T's doesn't appear to be polymorphic, but can't confirm them from either direction. Don't need to re-do because 443 reaches T's from bottom strand.
	2452	done - no changes
	443	done - POLY - bp 76858, A to G transition in DYS129 - 6 bp's away from CG5 exon 20 (7-3a3) intron/exon boundary!! Genpoly 130.
26		
27		

REDACTED

pcrCOSMID - DYS129



REDACTED

EXHIBIT 3

REDACTED

Gen Poly 130
- develop + read!

Results: [redacted]

odd 1	12403 -	C
odd 2	14480 -	C/T
odd 3	08539 -	C/T
	5110 -	T
	24267 -	T
	14937 -	T
	22688 -	T
NC	I-10606 -	T
	I-7710 -	T
	I-1703 -	T
	Ax20 -	T
	22825 -	?

Need to do same pcr on more FD + NC!!
also, other odd 1, odd 2, Shoft.
+ cosmid, etc.

→ See loading sheets 4/13/00

PCR O/N!!!

Northern

M - Label 102909 x 8
S - Label CG-4-Exon 1 x 8

Probe Neo P.N., Std H, MTN, + 2 F's in Northern Chuck
+ Formamide

Clean probes + hyb O/N.

REDACTED

SS

EXHIBIT 4

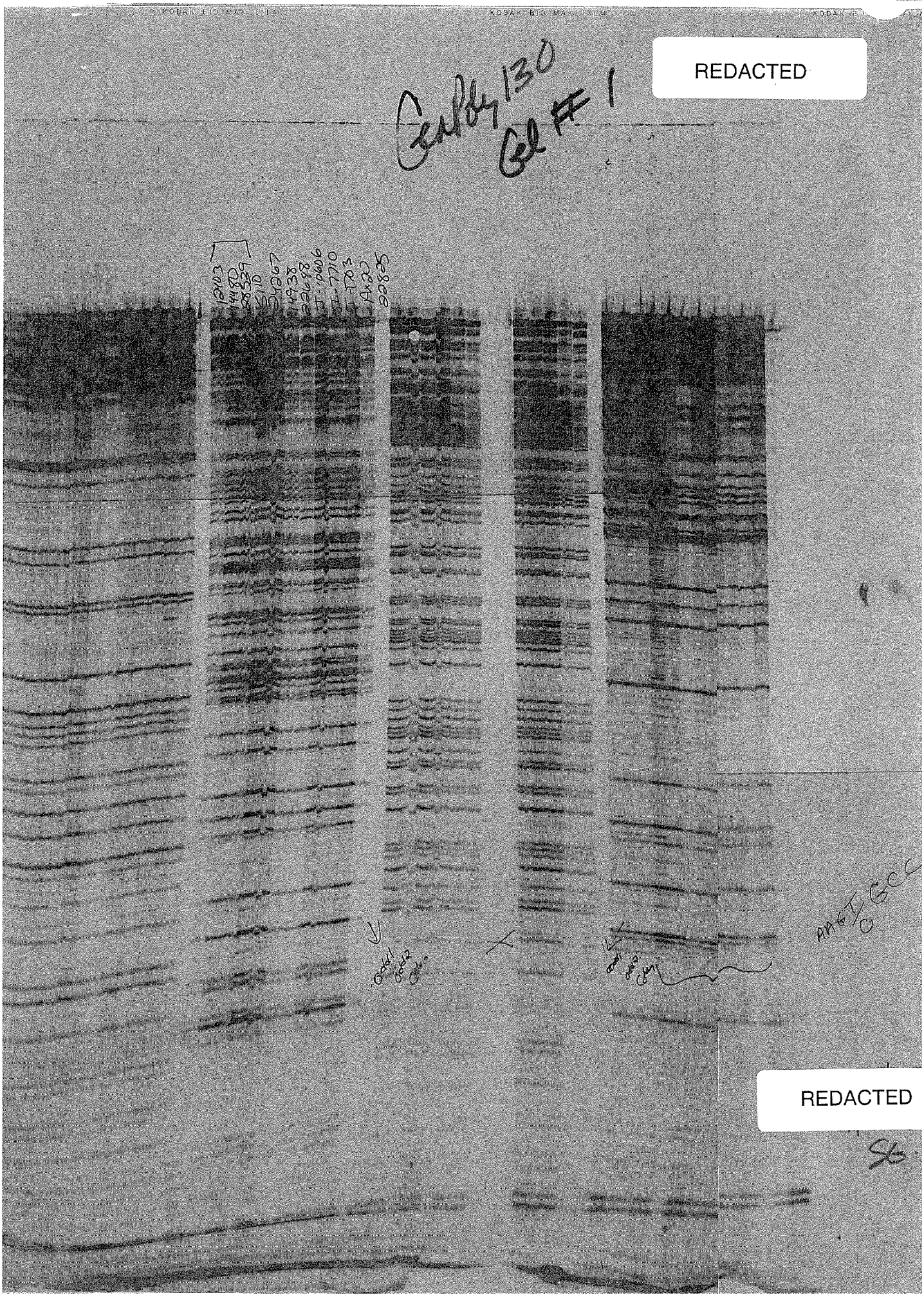

REDACTED
REDACTED

EXHIBIT 5

REDACTED

INTRON/EXON BOUNDRY SCREENING OF CG5

CYCLE SEQ OF EXON 19 300 B.P.

①	4390	GEN 473 #2007
②	18857	
③	14480	
④	12819	
⑤	COHEN	
⑥	AX20	
⑦	4390	CG5EXON 19R #2532
⑧	18857	
⑨	14480	
⑩	12819	
⑪	COHEN	
⑫	AX20	

GEN POLY 190

MIX

5 ul	DNA
1 ul	PRIMER
1 ul	33P
4 ul	ENZ + BUF
19 ul	H_2O
30 ul	TOTAL

94°	4 MIN	
94°	45 SEC	30 CYCLES
56	45 SEC	
72°	45 SEC	
15°	SOAK	

EXHIBIT 6

REDACTED

CG5 EXON 19 INTRON/EXON BOUNDRY

REDACTED

① 4390 - ST. AFF
②③④ 18857, 14480, 12819 } ODD 2
⑤ COHEN - ODD 3
⑥ AX20 - CONTROL

94° 4 MIN
94° 45 SEC
56° 45 SEC } 30 CYCLES
72° 45 SEC
15° SOAK

GEN 473
2007

CG5EXON 19 RL
2532